SUSTAINABLE CAPACITY PLANNING FOR HOSPITAL OPERATING THEATRE

LI ANG

(B.Eng. (Hons.), NUS)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF ENGINEERING

DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2009

Acknowledgement

I would most sincerely thank my supervisors, Dr Yap Chee Meng and Associate Professor Lee Loo Hay from the Department of Industrial and Systems Engineering, NUS, for their continual support and advice on my research project. They have not only introduced new concepts in searching algorithms and in modeling techniques to me, but they have also taught me how to realistically apply what I have learnt to the real world problem, as well as how to interact with people from the industry.

I would also like to thank Ms. Judy Tan and Ms. Luna Lee from the local hospital we have worked with for their enthusiastic sharing of their operational problem and for their generous provision of the data required for modeling and solving the problem.

Last but not the least, I would thank Ms. Chua Siang Li and Mr. Leow Kiok Liang from the National Health Group for their participation and provision of ideas as the research project was conducted.

Table of Contents

Acknowledgement **(i)**

Table of Contents **(ii)**

Summary **(iv)**

List of Tables **(v)**

List of Figures **(vi)**

Chapter 1 **Introduction……………………………….……………………...1**

Chapter 2 **Literature Review..…………………………………………......10**

2.1 Block-booking vs. FCFS - Implication on OT Decision Making Process ……………………..…………………………………...…..10

2.2 Recent OT Sequencing Studies…………………... ………………...…12

2.3 Recent OT Planning/Scheduling Studies………………………..........17

2.4 Our OT Capacity Planning Study……………………………………....23

Chapter 3 **OT Capacity Planning - Problem Description…..…………….....27**

3.1 Operational Characteristics……………………………….……......27

3.2 Current Practice and Motivation for Capacity Planning………....…..32

Chapter 4 **Simulation Model……………………………………………..37**

4.1 Scale of Simulation Model……………………..……………………....38

4.2 Data Preparation…………………………...…….……………….......40

4.3 Simulation Model Structure and Dynamics……………………………...43

4.4 Simulation Model Assumptions………………………………………..47

Chapter 5 **Model and Optimization Algorithm……………………….......49**

5.1 Problem Formulation…………………………………………………...49

5.2 Optimization Algorithm……………………………………………….......52

5.2.1 Algorithm Phase One – Generation of Departmental OT Plans

5.2.2 Algorithm Phase Two – IP Model

Chapter 6 Algorithm Implementation and Numerical Results…………58

6.1 System Parameters………………………………….………………...58

6.2 Phase One Algorithm Implementation and Results…….……………..60

6.3 Phase Two Implementation and Results……...………………………..68

Chapter 7 Conclusion ……………………………………….….……....74

List of References……………………......................……………….……..77

Appendix…………………………………......................………………79

Summary

The aging population and the limited governmental funding for healthcare are expected to overwhelm the healthcare systems in most countries in the forthcoming decades. It is of paramount importance for healthcare practitioners to learn to operate in such an unprecedentedly highly constrained environment. In general, adapting to their new roles involves the hospitals to build up a strong-enough financial wellbeing to allow for their continuous growth on the one hand; and on the other hand, to keep serving as many patients as possible in fulfilling their expected social roles.

The operating theatre is one of the most expensive resources within the hospital infrastructure. In this study, we consider the optimal utilization of the operating theatre resource to best address the expectations of the three major stake-holders involved in the hospital's surgery operations. Namely, these three parties are the patients, the surgeons, and the hospital itself. We first argue that the "block-booking" scheduling approach generally suits the need of the surgeons to have control over their own schedule, and of the patients to have transparency and certainty of their surgery time and date. Based on the block-booking framework, we propose an algorithm that best allocates the OT time to the surgeons from the various clinical departments. In particular, our approach combines the simulation methodology and MIP, and it bears with two conflicting objectives of both the profit and the patient throughput that are to be maximized. We further argue that this capacity allocation approach, if continuously applied, can help the hospitals to maintain fairness in the resource allocation on the one hand, and to maximally achieve their financial and social goals on the other.

List of Tables

Page

Table 2.1: Comparison between Block-booking and FCFS OT Systems with Respect to the OT Decision Making Processes……………………..…....11

Table 2.2: Comparison between Ozkarahan (2000) and Jebali et al (2006)…………………………………………………………...…..…15

Table 2.3: FCFS vs. Block-booking - Compared for the Three Major OT Stakeholders……………………………………………………….….24

Table 6.1: System Parameters for the Departments……………………………......59

Table 6.2: Study Period Information for Department #6 Surgeons………….....61

Table 6.3: Pareto Solutions with Performances at all Capacity Levels – Department #6……….………………………………………………..63

Table 6.4: Summary of Phase One Results – Number of Candidate OT Plans…………………………………………………………..….68

Table 6.5: Global Optimal Pareto Set…………………………………………..69

Table 6.6: Robustness of Departmental OT Plans Picked……………………..…72

List of Figures

Page

Figure 1.1: Change of World Population Made-up: 1950-2050…………………..…..1

Figure 1.2: "Break-even" between two Conflicting Objectives of a Hospital…….4

Figure 1.3: Operations Schedule under Block-booking OT System………………6

Figure 1.4: Dynamics of FCFS OT System…...…………………………………….7

Figure 3.1: Example of the Weekly OT Schedule…………………………………29

Figure 3.2: General OT Workflow for Elective Surgeries………………………..30

Figure 4.1: Decomposition of a Full OT Plan into its Elements………………….38

Figure 4.2: General Simulation Model Structure………………………………....43

Figure 4.3: Surgery Booking/Execution Process Dynamics…………………...…45

Figure 5.1: Major Mechanism in Phase One Searching Algorithm…………......56

Figure 6.1: Plotted Pareto Solutions - Department #6……………………………65

Figure 6.2: Plotted Pareto Solutions- Department #1…………………………….67

Figure 6.3: Plotted Global Pareto Solutions……………………………………....70

Chapter 1: Introduction

The human race has run into an era when the population is getting old faster than ever before. It is projected that by year 2050, 21% of the population on our planet will be over 60 years old. This proportion, as reported for years 1950 and 2000, was only 8% and 10% respectively (UN, 2002). Figure 1.1 illustrates the tremendous change on the age made-up of the world's population from the mid-twentieth to the mid-twenty-first century.

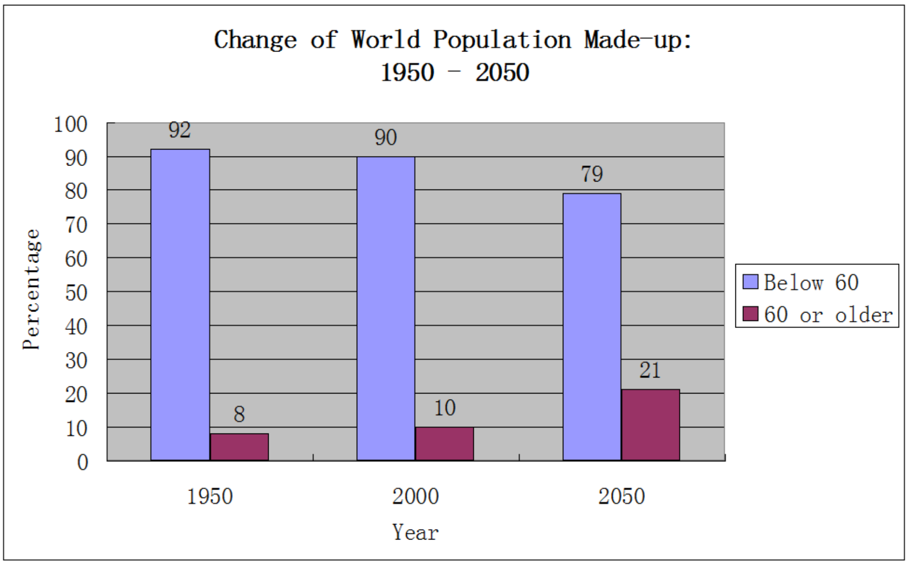

Figure 1.1: Change of World Population Made-up: 1950-2050

The aging population results in major consequences and implications for all facets of human life. In the healthcare systems for instance, the aging population will imply diminished workforce in the societies, lesser income tax collected by the governments; and consequently, higher burden for the governments to subsidize the medical treatments of their people on the per capita basis. Indeed, in a lot of developing countries where the rate of population aging is atrocious, whether the growth of the countries' economy is sufficient and sustainable enough for handling their aging people's healthcare has been attended with much awareness (UN, 2002).

In some developed countries on the other hand, the aging population has already started to create problems for the health systems. In the United States for example, the number of public hospitals (i.e. hospitals that are wholly owned and operated by the government) has declined 27% (from 134 to 98) from 1996 to 2002 in the major suburbs. It is thought that the increase in the uninsured population, which in fact has a lot to do with those non-working aged senior citizens, has drained public hospitals to near bankruptcy (Higgins, 2005).

In short, today's healthcare industry is characterized by the ever increasing patient loads and the (relatively) ever decreasing governmental subsidies. Operating in such a highly-constrained environment, the hospital administrators are required to deliberately search for ways that best utilizes their medical resources to achieve the optimal operational efficiency, and consequently the optimal medical outputs.

To achieve this goal, in our opinion, the first step will be that the hospitals rationalize the roles they "wish to play" and the roles they are "expected to play", in the context of the entire social system. A rather strategic-level clarification as such shall serve as the hospitals’ guide throughout their activities on the downstream tactical and operational levels.

We consider the roles concerning hospitals in the social system as two-fold. Specifically, on the one hand, as hospitals also operate in the form of “enterprises” that participate in the free-market economic system, they are required to strive to build up sustainable financial growths for themselves. Such growths will enable the hospitals to continuously recruit and train their people, update their infrastructure with latest technologies and equipments, and invest on the service innovation and process optimization projects. All of these will in turn help the hospitals to maintain a stable provision of healthcare to the society.

From the perspective of the public on the other hand, hospitals are viewed a lot like the safety nets for keeping *all* people’s lives. Entrapped in such a role, hospitals are expected to provide treatments as much as they can afford, and are expected to save as many lives as possible, regardless of the class, the race nor the wealth of the people.

It is not surprising to see that the role the hospitals “wish” to play for their own goods and the one they are expected of by the society are conflicting in nature. The social expectation, if completely obeyed, could drag the hospitals into bankruptcy long before the managements see any big number of “annual” reports. Yet on the reverse, if hospitals operate solely “profit-orientated” they may leave large numbers of low-income patients unattended, the consequence of which may be fierce social conflicts and crisis.

In our opinion, a rational hospital's administrators should carefully seek a balance between these two conflicting objectives they seek. Certain break-even point on the scale linking the two "objective poles" is desirable on which both satisfactory level of financial growth is preserved for the hospital, and appropriate amount of patients are served (see Figure 1.2). The hospitals have to carry out this balancing task when they are faced with any strategic planning problem for their operations.

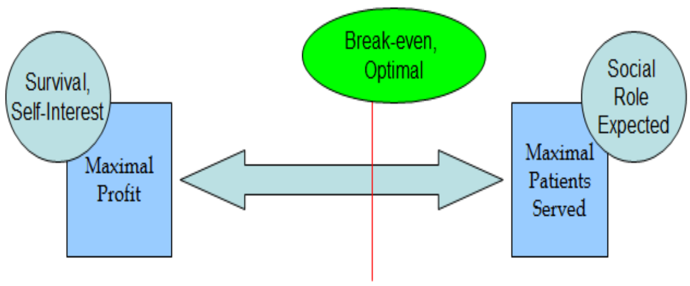

Figure 1.2: "Break-even" between two Conflicting Objectives of a Hospital

The operating theatre ("OT") is considered one of the most expensive resources within the hospital's infrastructure (Singapore National Health Group, 2005). Building and upgrading the OT facilities could easily involve millions of dollars; and largely because of this costly nature, the planning and scheduling of the OT activities has received attention from the research society ever since the early 1970's. A review of the recent OT planning and scheduling studies (since year 2000) will be provided in Chapter 2.

An OT may consist of several operating rooms ("OR"s). In general, surgery cases that are performed in the OT can be categorized into two groups, namely, the "elective surgeries" (i.e. cases that can be planned beforehand), and the "emergency surgeries" (i.e. urgent cases that have to be handled on the spot upon their appearances). Generally speaking, the emergency surgeries are handled by reserving certain ORs in the OT at all times. And as far as the elective surgeries are concerned, currently there exist two major frameworks under which the OT is deployed to serve them. These two frameworks are the "block-booking", and the "non-block-booking" OT systems.

In the block-booking system, the OT's capacity (characterized oftentimes by the one 8-hr or two 4-hr time blocks of each OR on each day in a week) is reserved purely for the usage of one particular surgeon or one particular group of surgeons in a clinical department. When a new request for elective surgery appears (either from an outpatient visit or an inpatient follow-up examination), the surgeon will book the case in the computer system with an estimated duration s/he will need for the case. After this, if no further change is required, both the surgeon and the patient will wait until the booked surgery date for the operation. Usually in a block-booking system, the un-booked capacity will be freed up two to three days before the actual surgery dates. This is to say that from the reservation "cut-off" time onwards, all the surgeons will be able to book their cases into the unused OT slots. Figure 1.3 shows an example of the schedule of operations under the block-booking OT system. Note that the shaded portion in each box represents the booked surgery duration in that particular OT time block. Also note that the system can arrange a surgery to be performed a few weeks from the time it arrives, depending on the availability of the department/surgeon's OT time.

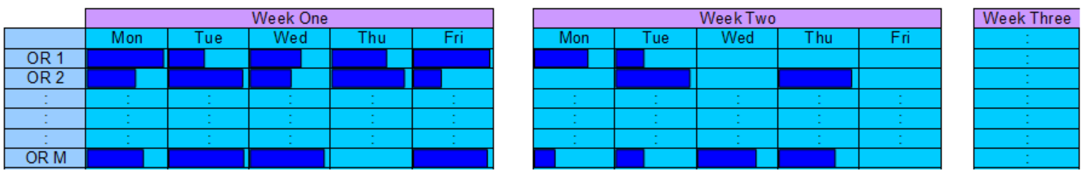

Figure 1.3: Operations Schedule under Block-booking OT System

The non-block-booking OT system is also called the FCFS (First-Come-First-Served) system. In such a system, instead of a confirmed weekly OT capacity plan, a list of "candidate surgeries" is carried over time that includes surgery requests from all departments. Whenever a new elective surgery request comes, instead of being directly booked in the computer system, the case will be informed to the OT manager, who will then assign a "tentative" surgery date for the case, based on the availability of the total OT capacity in the following days/weeks to his understanding. This tentative surgery date is then informed to the surgeon and the patient to prepare for the operation. At the end of every week, normally, a scheduling and sequencing program is run that reconfirms the actual cases to be done in the following week together with their exact sequences and their OR assignments. In building such schedules, the initial "tentative surgery dates" will be attempted to be kept, but changes to them also happen frequently due to all sorts of unforeseen conditions. Figure 1.4 illustrates the dynamics of a FCFS OT system.

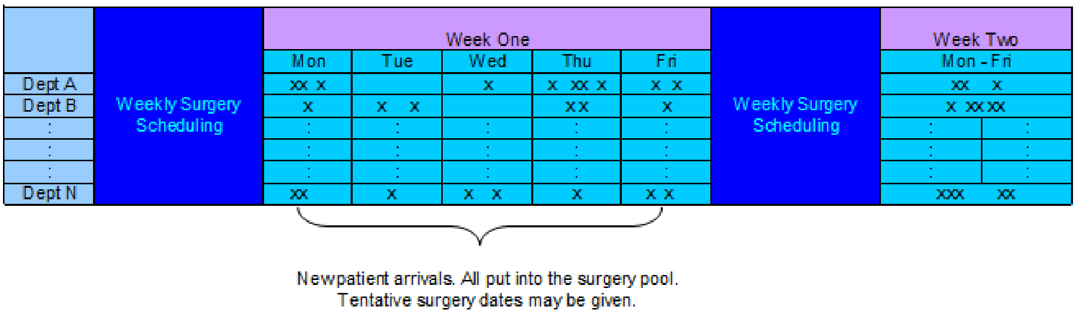

Figure 1.4: Dynamics of FCFS OT System

In this thesis, we consider the approach to the optimal utilization of the OT capacity. We first carry out a comparison between the block-booking and FCFS OT systems in term of the preferences of the three major stakeholders on the OT activities, namely, the surgeons, and patients, and the hospital. We conclude that the block-booking system in general suits the needs of the surgeons in having control over their own schedule, and of the patients in having transparency and certainty of their surgery date/time.

Based on the block-booking system, we further propose an approach that allocates the OT capacity to surgeons coming from various clinical departments on a weekly basis. This actual study is done in collaboration with a local not-for-profit hospital in Singapore. The objective of the capacity allocation is to simultaneously achieve the maximal profit and the maximal patient throughput, which are in accordance to the two objectives the hospitals rationally ought to seek as we discussed earlier.

In solving the capacity allocation problem, we propose an algorithm that combines the advantages of simulation methodology and Integer Programming (IP). In particular, the algorithm consists of two phases due to the problem-specific structure of the system where surgeons are grouped by their clinical departments. In our algorithm, the two phases respectively seeks to identify the optimal “departmental”, and the optimal “hospital-wide” OT capacity plans. The simulation model is developed to mimic the procedures in which surgery requests are received, handled, and ultimately served or cancelled in the hospital, based on which the performance of any OT plan is evaluated. The IP model serve to synthesize the optimal departmental OT plans as generated in the Phase One, and delivers the “global” optimal OT plans for the hospital.

The rest of this thesis is organized as follows. In the next chapter, we will present a literature review of the recent studies concerning the management of OT activities. We will highlight the problems being addressed under both the block-booking and FCFS OT systems, and we will discuss in details the pros and cons of these two systems from the perspectives of the surgeons, the patients, and the hospital. In Chapter 3 we describe our OT capacity allocation problem. This includes detailed operational characteristics of our OT system, and our original motivation to carry out this optimization study. In Chapters 4 and 5 we will explain the two-phase algorithm that combines simulation and IP to identify the optimal OT plans. Specifically, the setup, data preparation, and the detailed structure and dynamics of the simulation model will be presented in Chapter 4, and the two-phase algorithm will be explained in details in Chapter 5. In Chapter 6 we illustrate the implementation of the algorithm and show the numerical results as we apply the algorithm to solve our problem. Chapter 7 will conclude our study.

Chapter 2: Literature Review

This chapter presents a review of the latest studies that address the OT capacity management problems. In fact, in practice the management of OT activities often consists of two levels of decision-making. Specifically, there is a higher "planning/scheduling" level, on which the surgeries to be performed in the forthcoming time period (usually the following week) are determined; and below this, there is a "sequencing" level, where the exact sequences and room assignments are decided for the cases that will be performed on the very next day. We present in this chapter the recent studies that address problems on both of these levels.

After these, we present a comparison between the block-booking and the FCFS OT systems from the perspectives of the three major OT stakeholders, i.e. the surgeons, the patients, and the hospital itself, with the conclusion that the block-booking system in general suits more of the needs of these three parties. We then discuss our own research problem, which is the allocation of OT capacity under the block-booking system. In doing so, we have reexamined the recent block-booking planning/ scheduling studies for their relevance to our problem.

2.1 Block-booking vs. FCFS - Implication on OT Decision Making Process:

Recall our discussion in Chapter 1, the block-booking and the FCFS OT systems are the two main-stream frameworks under which the OT is deployed to serve the patients. Despite the rather distinctive operational characteristics these two systems have, essentially, each system can be viewed as a set of “rules” imposed on how the elective surgeries are accepted, handled, and served in the OT system continuously over time. In other words, each of these two OT systems has within itself a specific approach to arranging the surgeries to be performed in the forthcoming time periods, and thus a specific theme on the “planning/scheduling” level of the OT activities.

Table 2.1 presents a comparison between the block-booking and the FCFS OT systems on their respective decision making processes for the OT activities.

Table 2.1: Comparison between Block-booking and FCFS OT Systems
with Respect to the OT Decision Making Processes

<table>
<tr><th>System</th><th>Block-booking</th><th>FCFS</th></tr>
<tr><td rowspan="3">Planning/
Scheduling</td><td>Slot Allocation</td><td rowspan="3">Weekly surgery scheduling from "case pool"</td></tr>
<tr><td>Slot Assignment</td></tr>
<tr><td>Routine Workflow</td></tr>
<tr><td>Sequencing</td><td colspan="2">Daily decisions on sequences and OR assignments for next-day surgeries</td></tr>
</table>

In the FCFS OT system, the procedure is relatively straight-forward. At the beginning of every week, a scheduling task is carried out that picks out (from the “waiting list” of the surgeries) the cases to be performed and packs them into each OR on each day in the week. Afterwards, on every day before the operations, a sequencing of the planned surgeries is performed that reconfirms the availability of the surgeons, their assisting panels and the medical equipments, and that puts the surgeries into exact sequences in each OR.

In a block-booking OT system, on the contrary, the practice consists of a few more elements. First of all, a capacity planning task has to be carried out that allocates the appropriate amount of OT slots to the various surgeons/surgeon groups demanding them. After this, these allocated OT slots have to be assigned to the actual ORs and on the actual days in a week according to the equipments' availabilities and to surgeons' preferences. When both of these two steps are done, routine practice of surgery booking and execution can be performed in the block-booking OT system. It needs to be noted that same as in the FCFS system, in the practice of the block-booking system there also has to be the stage of "sequencing" before each day's surgeries.

2.2 Recent OT Sequencing Studies:

As we have discussed, before carrying out the operations on each surgery day it is necessary to re-confirm the surgeries and to sequence them out in the various ORs. The purpose of doing this, essentially, is to ensure a smooth patient flow in the OT, and to reduce the risk of over- and under- running the very costly OT facility.

It needs to be noted that in practice, the planned and scheduled surgeries can be cancelled due to a variety of reasons. Some of these reasons include: 1) surgeon's perception that the case is unnecessary or too dangerous to be performed, 2) patient's increased fear of the surgery, and 3) patient's decision to refer to another hospital for the surgery. As such, in order to ensure the high utilization rate of the OT resources it is the norm that OT managers tries to schedule or book more cases than the OT's actual capacity in the up-stream planning/scheduling level. However, the downside of this is the risk of having an overloaded schedule for every surgery day, and thus the need to again "pick out" the most important cases before each day during the sequencing. The cases that are left over normally will receive highest priority to be served within the next couple of days.

In the recent literature, Ozkarahan (2000) and Jebali et al (2006) present rather sophisticated approaches in conducting the surgery sequencing task. Specifically, Ozkarahan (2000) uses goal-programming approach to simultaneously achieve several objectives in building the surgery sequences and room assignments. Issues and factors considered in his study include: a) the available duration of each OR on the surgery day, b) the compatibility between the ORs and surgeries from the different clinical departments, c) the availability of the intensive care unit (ICU) beds, d) the priority for each case to be conducted on the surgery day (as opposed to being delayed to the next days), and e) the pre-planned OT time for each clinical department. Surgeries are packed into the various ORs so as to achieve goals related to each of the issues/factors listed above (refer to Table 2.2 for the actual goals modeled in the study). Afterwards, the sequence of surgeries in each OR is arbitrarily determined that realizes the optimal solution of the goal-programming model.

Jebali et al (2006)'s study is a lot similar to Ozkarahan (2000)'s in that they also model the sequencing problem as one that packs the surgeries into the ORs, and that the factors they consider also cover: a) the available duration of each OR, b) the compatibility of the ORs to the surgeries, and c) the availability of the ICU beds. The difference with this study is that the modeling approach adopted is the Mix-Integer-Programming (MIP), and that the latter two of the three concerns above are modeled as constraints that must be satisfied. An additional constraint is the available time of the surgeons. The objective of the study is to simultaneously minimize the cost of either over- or under-utilizing the OT as well as the cost of delaying the surgeries to the next days. Note that this objective is precisely the first and the last goals in Ozkarahan (2000)'s goal-programming model.

After packing the surgeries into the different ORs, Jebali et al (2006) develop an extra MIP model to derive the exact sequences of surgeries in each OR. The constraints of this model make sure that the surgeries are scheduled closely in line with the available hours of the surgeons and the ICU beds throughout the surgery day. In essence, this extra step results in higher "resolution" towards the surgery sequences as compared to Ozkarahan (2000).

Table 2.2 presents a comparison between Ozkarahan (2000) and Jebali et al (2006)’s surgery sequencing studies.

Table 2.2: Comparison between Ozkarahan (2000) and Jebali et al (2006)

<table>
<tr><th colspan="2">Study</th><th>Ozkarahan (2000)</th><th>Jebali et al (2006)</th></tr>
<tr><td rowspan="9">Packing</td><td rowspan="5">Objective</td><td>a) Minimize over- and/or under-utilization of the OT</td><td>a) Minimize over- and/or under-utilization of the OT</td></tr>
<tr><td>b) Minimize incompatibility of surgeries in the ORs</td><td rowspan="3">b) Minimize delay of surgeries</td></tr>
<tr><td>c) Minimize exceeding ICU bed availability</td></tr>
<tr><td>d) Minimize delay of surgeries</td></tr>
<tr><td>e) Maximize realization of pre-planned OT time for each department</td><td></td></tr>
<tr><td rowspan="3">Constraints</td><td rowspan="3">None</td><td>a) Compatibility of surgeries in the ORs</td></tr>
<tr><td>b) Availability of ICU beds</td></tr>
<tr><td>c) Availability of surgeons</td></tr>
<tr><td>Approach</td><td>Goal Programming</td><td>Mixed-Integer-Programming</td></tr>
<tr><td rowspan="3">Sorting</td><td>Objective</td><td rowspan="3">Arbitrarily done</td><td>Minimize over-running of OT</td></tr>
<tr><td>Constraints</td><td>Available hours of surgeons and ICU beds</td></tr>
<tr><td>Approach</td><td>Mixed-Integer-Programming</td></tr>
</table>

Besides Ozkarahan (2000) and Jebali et al (2006)’s studies, in the recent literature, Hsu et al (2003) consider sequencing the surgeries in an ambulatory surgical center. Surgery procedures in an ambulatory surgical centre have a lot less complexity as compared to those in an operating theatre in that there exists no over-booked surgery list, and that the cases can be performed by any surgeons due to their simplicities. The objective of the study is to shorten the make-span as well as to reduce the number of nurses required to look after the post-surgery patients, by means of building a "smooth" patient schedule. The method used is TABU search. In our view, this particular approach could hardly be generalized to surgery sequencings in the ordinary OT settings, mostly because of the simplicity of the system under study. In addition, Dexter et al (2002) evaluate a few heuristics in surgery sequencing with respect to the over-utilization rate of the OT.

Despite the capability of the aforementioned studies to pick out the important surgeries from the daily case pool and to arrange them into the appropriate ORs, one major shortcoming of their methodologies is that they all assume "deterministic" lengths for the surgery times. In practice, the actual time taken for a surgery can be very different from what is expected during the planning, and can be highly random (Strum et al 2000a and 2000b). Due to the random surgery times, it often happens in reality that scheduled surgeries get badly delayed or even cancelled because of the surgeries before them. In general, delays and cancellations of surgeries result in major frustrations among the surgeons and patients, who both are important stakeholders in the OT system.

In order to alleviate this problem, some researchers propose to give in the utilization of the OT and schedule a gap between two surgeons' cases in order to ensure the reliability of second cases starting time (Dexter et al, 2001).

Dealing with the same problem, Marcon et al (2001) present an analytical model that incorporates the randomness surgery duration issue into the surgery sequencing. In the model, the surgery durations are assumed to be randomly distributed with known distributions. Similar to the studies of Ozkarahan (2000) and Jebali et al (2006), Marcon et al (2001)'s model also tries to pack the candidate surgeries into the given OR available hours, but with the objective of the maximizing the "probability" of finishing all the scheduled cases on-time. The model is solved by an analytical approach proposed by the authors.

We consider Marcon et al (2001)'s study a good starting point of deepened and more realistic research on surgery sequencing. However, more basic issues such as the availability of surgeons and resources e.g. ICU beds and certain particular equipments have yet to be incorporated into the model. This gap ought to be filled by future researchers.

2.3 Recent OT Planning/Scheduling Studies:

As discussed in Section 2.1, the block-booking and FCFS OT systems have inherently different approaches in planning and scheduling the surgeries over the weeks.

In fact, the surgery planning/scheduling in a FCFS OT system shares great amount of similarity with its down-stream surgery sequencing activity which is conducted on the daily basis. Specifically, before starting the operations in each week, the OT manger running a FCFS system will have to pick out, from among the waiting-list of surgeries, the cases that will be performed in the week, and assign them into each OR on each day in the week. And this task is done by considering similar factors as in the surgery sequencing such as the availability of the surgeons and the other resources (e.g. ICU beds, supporting staff, etc), the equipment compatibility between the surgeries and the ORs, and so forth.

However, one additional character with the surgery planning/scheduling in the FCFS system as to the surgery sequencing is the consideration of "hospitalization" and/or "deadline" dates of the cases. In practice, at the time the weekly OT schedule is formed, some patients might not have been admitted into the hospital yet. For such patients, what are available are their planned admission (i.e. "hospitalization") dates, and sometimes the "deadline" dates set for the surgeries. Naturally, in scheduling the surgeries, all cases have to be arranged between their respective hospitalization and deadline dates.

Basically, the scheduling of surgeries in a FCFS OT system tries to get all patients operated ASAP after they are admitted into the hospital. This objective is sought in order to reduce the resources (e.g. inpatient beds, nurses, etc) consumed in taking care of the patients before their operations, and therefore, to reduce waste.

In the recent literature, Guinet and Chaabane (2003) present a surgery scheduling study that incorporates considerations of a full range of factors that cover: a) availability and daily capacity of each surgeon, b) availability of each OR on each day, in terms of both “regular hours” and “allowed over-time hours” with extra operating costs, c) compatibility of surgery types and the specific ORs as well as d) the hospitalization and deadline dates of the cases. The objective of the study is to find the optimal OT schedule with the minimum cost of over-time operations and the cost of delaying the surgeries after the patients’ admissions. An MIP model is built for the problem and a prime-dual heuristic is developed to solve the model.

On the other hand, Lamiri et al (2008) study a surgery scheduling problem with the same objective as Guinet and Chaabane (2003)’s, i.e. to minimize cost of over-running the OT and delaying surgeries after admissions. However, no availability and compatibility constraints are tackled in the model. The study also lends consideration of emergency demands of the OT, which appear only on the surgery day yet have to be served on the spot upon their arrivals. A hybrid approach that combines computer simulation and the MIP model is developed to solve the problem.

Similar to the case of surgery sequencing, one major shortcoming of the mainstream surgery planning/scheduling studies under the FCFS OT system is the deterministic assumption of surgery durations, and the rendered risk of under-utilizing the OT or surgery cancellations due to OT congestions. With this regard, Hans et al (2008) tackle the surgery scheduling problem in a stochastic framework. In their study, the mean and the variance of each surgery's duration are assumed to be known, and the model packs the surgeries into the ORs with the objective of maximizing the OT utilization and minimizing the risk of over-time, and thus cancelled patients. A variety of constructive heuristics and local search methods are presented to solve the planning model.

Despite the capability of approaching the rather stochastic nature of the real-world OT operations, Hans et al (2008)'s study inherently lacks consideration of all the resource availability and compatibility constraints, nor the cost issue of the delayed surgeries. Further research is still required to sew up these gaps.

Now refer to our Table 2.1, recall that in a block-booking OT system the planning/scheduling of surgeries initially involves the tasks of: 1) planning the total OT capacity (by allocating all the capacity blocks to all the surgeons); and 2) assigning the allocated blocks to the surgeons in the actual ORs and on the actual days in a week. After the assignment is done, straightforward routines are followed in booking and conducting the surgeries under each surgeon.

It therefore follows naturally that the focus of the surgery planning/scheduling studies in the block-booking OT systems lies on deriving the best strategy to allocate the OT capacity slots to the surgeons. The subsequent task of assigning the allocated slots to the actual ORs and actual days can be done through applying some simple assignment models, or even trivially done by hand.

In the recent literature, Blake and Donald (2001) and Blake et al (2002) tackle the allocation of the OT capacity slots together with their assignments, by using very similar approaches. The problem under consideration is the periodic revision of the OT capacity plan (which is conducted 3 to 4 times a year according to the authors) on the level of clinical departments. According to the authors, the need of such revisions arises due to a few practical issues such as the additions and/or reductions of available OT capacity over time, the incoming and leaving surgeons in certain departments, the expected rise and/or drop of patient demands, and so forth.

The objective of these two similar studies is to minimize the discrepancy on the number of OT slots granted to each department with respect to the previous OT capacity plan. A set of constraints are considered in the study, which include: a) upper and lower bounds on the total amount of slots each department can receive in the whole week, b) upper and lower bounds on the amount of slots each department can receive on each day (to ensure sufficient rest for surgeons in the department), and c) limits on the reduction of capacity slots in each department as compared to the previous plan. An IP model is developed to solve the problem.

Blake and Donald (2001) and Blake et al (2002)'s approach can effectively preserve the stability of the OT capacity plans over time, by seeking the minimal cross-plan discrepancies and by means of the various constraints set in the model. However, this approach inherently lacks the consideration of the "utilization efficiency" issue for the hospital. In particular, it is highly possible that departments whose surgeons work extremely inefficiently and/or whose surgeries incur large fiscal loss to the hospital keep receiving high levels of OT capacity in every revision simply because they were provided with it from the very beginning. In our opinion, more sensible performance metrics that directly reflect the hospitals' concerns (such as the operational efficiency and/or the profitability) should be used to evaluate the OT capacity plans.

With this regard, Kuo et al (2003) develop an LP model to maximize the general fee generated by the OT capacity plan. In their scenario, the number of capacity slots allocated to certain surgeons should be fixed; while for the other surgeons, these can vary between certain ranges. Different surgeons are expected to have different "rates" in bringing in "general fee" for the hospital, and thus, the LP model tries to pick out the best distribution of the slots that generate the highest level of "general fee", subject to the capacity boundaries constraints.

One limitation of this study, as we consider, is the assumption of linear relationship between the allocated capacity to the surgeons and the amount of "general fee" that will actually be generated. This assumption lacks consideration of the rather dynamic process how patients flow through the OT system, and the impact of the variability of the surgery durations and the surgery profits on the total profits made. On the other hand, as we have discussed in Chapter 1, focusing merely on the profitability side of the OT plans bears the risk of discouraging the less wealthy patients from having surgeries, which consequently could cause the failure of the hospital in fulfilling its social goals.

2.4 Our OT Capacity Planning Study:

In our study, we consider the strategic-level planning of the OT operations.

We first take a closer look into the block-booking and the FCFS OT systems, from the perspective of the important stakeholders involved in OT operations. This step will help us to gain insights to decide on “where” and on “how” we should step out on our study.

According to Lovejoy and Li (2002), there are three major parties involved in the OT operations whose preferences towards their experiences/outcomes at the OT collectively determine the overall success of OT management. These three parties include the patients, the surgeons, and the hospital itself.

In Table 2.3, we present a comparison between the block-booking and FCFS OT systems based on the experiences/outcomes of the three parties at the OT.

Table 2.3: FCFS vs. Block-booking
- Compared for the Three Major OT Stakeholders

Stakeholders	OT System Framework	
	FCFS	Block-Booking
Patients	Estimated surgery date given which is subject to change; patient passively accepts the estimated date.	Confirmed surgery date given on the spot of consultation; patient free to decide on changing hospital.
Surgeons	Surgery date not guaranteed; surgery starting time not known until the day before operation.	Specific surgery date/time under control; surgeons can arrange for other activities.
Hospital	Passively aligning surgeries with OT capacity, no strategy embedded in planning, yet complaints arise on lack of fairness.	With continuous revision of the schedule, fairness among surgeons is established, complaints reduced, and the hospital achieves optimal output in terms of profit and served patient load.

First of all, note that in a FCFS OT system, only “tentative” surgery dates are given out for the patients and the surgeons when doing the surgery booking. The actual surgery dates are subject to change due to a variety of unforeseeable circumstances; however, the patients will have to accept the estimated surgery dates upfront and make arrangements for their hospitalization, and the surgeons will have to cancel their other plans on the estimated surgery dates to prepare for these “unreliably-booked” cases. These make the situations of both the patients and the surgeons passive in the system.

On the contrary, the dynamics of the block-booking OT system inherently ensure both the patients and their surgeons the dates of surgery on the spot of surgery booking. Patients can freely choose to refer to other hospitals for surgery if s/he perceives the waiting time to be unbearably long, and meanwhile surgeons can arrange for other activities on those non-surgery days with confidence.

On top of the preferences of the patients and surgeons, it can be also seen that the FCFS OT system inherently lacks the flexibility for OT managers to inject any rules that can optimize the overall outcome of the process, e.g. patient load, profit, resource utilization, etc. Primarily, this is due to the fact that all the booked surgeries have to be served sooner or later, and that the essential role of the OT managers in such a FCFS system is merely to arrange for the schedules of the surgeries. However, within the block-booking OT framework, freedom is within the hands of OT managers to allocate different levels of OT capacity to the different surgeons, and eventually to achieve the most desired levels of outcomes. More importantly, under the block-booking OT system competition mechanisms can be introduced into the practice by which surgeons will proactively work with high efficiency in order to keep their granted OT capacity (if not to be qualified for higher levels). This in turn will pull up the overall performance of the entire OT.

Therefore in our study, we base our strategic-level OT planning on the framework of the block-booking OT system. We have worked with a local hospital whose default

OT practice is the block-booking system. We develop a more rational and more sophisticated approach as compared to Blake and Donald (2001), Blake et al (2002), and Kuo et al (2003)'s studies in allocating the OT capacity, through which we seek to achieve the optimal level of patient throughput, and the optimal level of profits for the hospital. We further claim that if our approach is continuously applied, then the hospital will be able to maximally achieve their financial and social goals whilst preserving fairness in the capacity allocation. From the next chapter onward, we will present in details the problem characteristics, the model, and the solution approach we develop to solve the OT capacity planning problem.

Chapter 3: OT Capacity Planning - Problem Description

In this chapter, we describe our block-booking OT capacity planning problem in detail. We first explore the operational characteristics of the OT system at the hospital we worked with. These include an overview of the OT, and a detailed workflow of how elective surgeries are handled and served in the system. The original capacity allocation method at the OT and the motivation for conducting the study are discussed in the later part of the chapter.

3.1 Operational Characteristics:

The local hospital we have worked with is a large-scale, acute-care, tertiary hospital that operates with a manpower of over 3,000 professional staff. Throughout its numerous clinical departments and specialized services centers, the hospital is able to provide a comprehensive line up of medical services to meet the requirements of both its international and local patients.

The hospital possesses two OTs, namely, the Major Operating Theatre (the "MOT") and the Day Surgery Operating Theatre (the "DSOT"). Among these two OTs, the MOT is located in the hospital's main building, and is used as the hospital's major facility to carry out surgeries of a wide spectrum of illness and seriousness. A few major clinical disciplines that use the MOT include Orthopedics, Cardiology, Neurology, Urology, and Pediatrics. The DSOT, on the other hand, is located in a separate building in the hospital, and is used mostly to handle the simpler "day-surgery" cases, i.e. minor surgeries for which patients are admitted, operated, and discharged from the hospital all on the same day. Examples of such cases include removal of skin lesions, hernia repairs, and most of the dental procedures.

The MOT consists of a total of 12 ORs. Among these, two are reserved for the emergency surgeries, and ten are used for elective surgeries. The DSOT has 6 ORs, and all of these are used for elective surgeries.

Both MOT and DSOT practice block-booking. Being more specific, both of these two OTs operate five days a week, and two sessions a day. The first daily session starts from 8:30am and ends at 12:30pm; and the second session starts at 1:30pm and ends at 5:30pm. Therefore, for each of the hospital's 16 elective ORs (10 from MOT and 6 from DSOT) there are 2 (sessions/day) * 5 (days/week) = 10 capacity slots in a week that are available for elective surgeries. Altogether, the hospital has a total of 160 slots/week OT capacity.

Upon request from the hospital, in this study we have considered only the capacity allocation at MOT. Without further clarifications, in the later text we will refer to MOT directly as the "OT".

As is the case in any block-booking OT system, the utilization of the hospital's OT is controlled by means of a weekly OT schedule, on which the usage of each OR in each session in a week is granted to one particular surgeon/surgeon group in the hospital. However, it needs to be highlighted that in this particular hospital we worked with, the OT slots are not shared between any surgeons in any clinical department. Each slot is given out purely as "dedicated" to a single surgeon. This practice is due to the fact that surgeons are perceived as extremely important resources by this hospital, and thus, they are required to be provided fully with their own time in planning and conducting their cases. As a matter of fact, it is quite common that some famous surgeons receive multiple OT slots in a week for their surgeries. An illustration of the weekly OT schedule in the hospital is shown in Figure 3.1. Note that some surgeons receive "full-day" capacities.

<table>
<tr><th></th><th>OR1</th><th>OR2</th><th>OR3</th><th>..</th><th>OR10</th></tr>
<tr><td>Mon</td><td>Orthopaedic Surgery</td><td>Orthopaedic Surgery</td><td>Urology</td><td>..</td><td>Cardiac</td></tr>
<tr><td>AM</td><td rowspan="2">Dr. Ang</td><td>Dr. Teng</td><td rowspan="2">Dr. Wang</td><td rowspan="2"></td><td rowspan="2">Dr. Lee</td></tr>
<tr><td>PM</td><td>Dr. Shao</td></tr>
<tr><td>Tue</td><td>Orthopaedic Surgery</td><td>Orthopaedic Surgery</td><td>Urology</td><td>..</td><td>Cardiac</td></tr>
<tr><td>AM</td><td rowspan="2">Dr. Lim</td><td rowspan="2">Dr. Hui</td><td rowspan="2">Dr. Zhang</td><td rowspan="2"></td><td rowspan="2">Dr. Chew</td></tr>
<tr><td>PM</td></tr>
<tr><td>Wed</td><td>Orthopaedic Surgery</td><td>Dental</td><td>Urology</td><td>..</td><td>Cardiac</td></tr>
<tr><td>AM</td><td>Dr. Poh</td><td rowspan="2">Dr. Low</td><td rowspan="2">Dr. Ng</td><td rowspan="2"></td><td rowspan="2">Dr. Yap</td></tr>
<tr><td>PM</td><td>Dr. Xie</td></tr>
<tr><td>Thur</td><td>Orthopaedic Surgery</td><td>Dental</td><td>Urology</td><td>..</td><td>Obstetrics & Gynaecology</td></tr>
<tr><td>AM</td><td rowspan="2">Dr. Pin</td><td>Dr. Leong</td><td rowspan="2">Dr. Chong</td><td rowspan="2"></td><td>Dr. Wong</td></tr>
<tr><td>PM</td><td>Dr. Chua</td><td>Dr. Hee</td></tr>
<tr><td>Fri</td><td>Orthopaedic Surgery</td><td>Dental</td><td>Urology</td><td>..</td><td>Obstetrics & Gynaecology</td></tr>
<tr><td>AM</td><td rowspan="2">Dr. Pho</td><td rowspan="2">Dr. Chan</td><td rowspan="2">Dr. Wong</td><td rowspan="2"></td><td rowspan="2">Dr. Chou</td></tr>
<tr><td>PM</td></tr>
</table>

Figure 3.1: Example of the Weekly OT Schedule

The basic procedure how an elective surgery is received, handled, and subsequently conducted in the hospital's OTs is summarized in Figure 3.2.

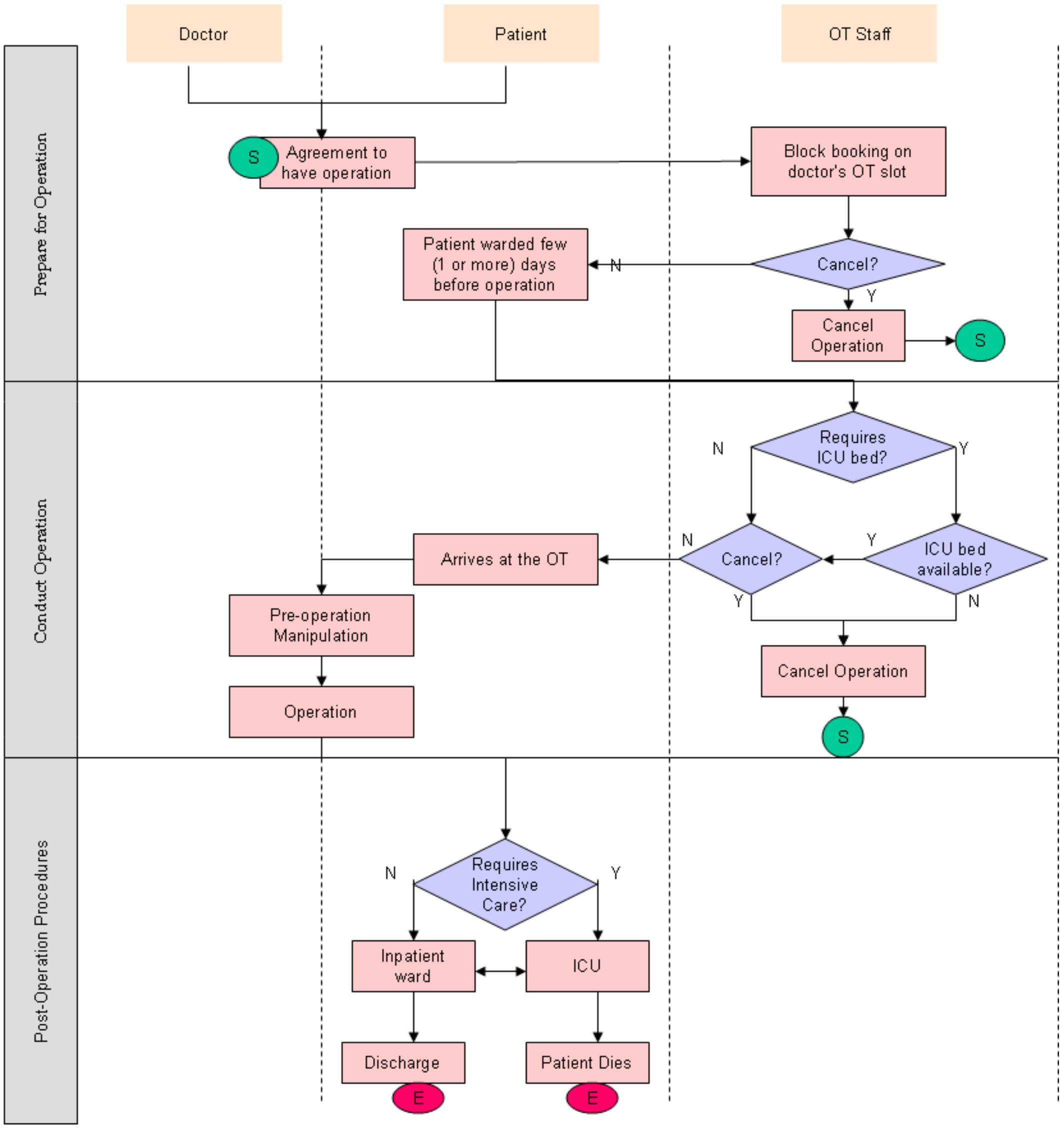

Figure 3.2: General OT Workflow for Elective Surgeries

Specifically, the demand for each elective surgery arises when a doctor and a patient make an agreement to have a surgery. In practice, this could happen during the patient's visit to one of the hospital's outpatient clinics, and a lot of times, in the inpatient wards when the doctors routinely check on the patients' conditions.

Upon making this surgery agreement, the OT staff will help the surgeon to book the case into the computer system. In particular, in doing this the surgeon is required to provide an estimated duration s/he needs for the case. Based on the estimation, the staff will look up the OT timetable, and find for the surgeon his/her next available week that has sufficient time for the case. If this time is acceptable for both the surgeon and the patient, then the case will be booked into the system. In some other cases, however, the patient may find the waiting time for the surgery to be too long, and s/he may choose to directly leave the hospital and refer to some other providers for the surgery. The surgery demand is thus lost in such a case.

Once a case is booked, the patient will start waiting for the operation. However, throughout the waiting, the case can be canceled due to a variety of reasons as we have mentioned earlier on. In some cases, the cancelled surgeries are sought to be booked again in the system; but most of the times they would leave the system and thus become lost demand. When a case is cancelled, the OT staff will eliminate the case from the computer system, and the time slot that is freed up can be used for booking another surgery. If the case is not cancelled, then the patient will be warded one or a few days before the operations date.

On the surgery date, the majority of patients will follow a standard routine to get anesthesia, and then operated. However, if an ICU bed is required after the surgery, then the surgeon needs to check with the ICU staff early in the morning for the availability of bed. If no ICU bed is available, the case will have to be canceled on the spot and get rescheduled at a later time.

After the operation, depending on the condition the patient will either be warded or sent to the ICU for further observation and treatments. In case the patient is sent to the ICU, if there is a post-operations complication, then s/he may be sent back to the OT for a follow-up procedure. At other times, s/he will be sent to the ward after the condition become stable. There are also times, unfortunately, when patient dies in the ICU or even during the surgery.

3.2 Current Practice and Motivation for Capacity Planning:

Conventionally, the hospital deploys a two-stage approach to allocate the OT capacity slots to its surgeons. At first, the central hospital management decides, on an aggregate level, the total number of slots to provide for each of the clinical departments that involve elective surgeries. There are ten such departments in the hospital in total. After this centralized capacity allocation is done, each clinical department will decide on its own how to distribute its received capacity slots to its surgeons.

Once decisions of slots allocation are made within all the departments, the exact weekly OT schedule of the surgeons will be formed. This task is performed by the OT staff through collectively gathering the preferred time of surgery from all the surgeons, and then constructing the actual schedule by putting these together.

The allocation of the OT slots has been rather rigid in the hospital, with revisions to it made only over long periods of time.

Before our study, the OT capacity allocation in the hospital, both on the centralized "hospital" level and on the "departmental" level, are decided based on experience. According to the hospital, among a few other factors and concerns, throughout revising the slot allocation the "utilization" of the allocated OT time is used as the primary performance indicator and benchmark for decision making. More specifically, in every revision of the capacity distribution, the hospital will collect data pertaining the duration that each department and each surgeon spent in the OT in the past period, and the utilization rate of the allocated OT slots is then calculated by applying the formula:

$$\text{Utilization} = \frac{\text{Average Weekly OT Usage over the Past Period}}{\text{Allocated Weekly OT Time}}$$

Particularly, in the formula, the OT usage includes both the time actually spent on conducting the surgeries, i.e. the surgery time, and that spent on setting up the equipments and cleaning the OR before and after the surgeries, i.e. set-up/clearance time. Both the nominator and denominator quantities are measured in minutes.

After the calculation, adjustments on the capacity allocation are made. This is done firstly by the hospital re-assigning a few slots from the department(s) that has relatively lower utilization rates to those with higher. And later, within each department, the administrators will re-arrange the slots from the rather "idler" surgeons to those "busier" ones.

According to the hospital management, the rational behind this "utilization-based" capacity allocation approach is to create the sense of "fairness" among the various departments and their surgeons (i.e. the more work performed, the more capacity given); and hopefully, to increase the patient throughput by allocating more capacity to the departments and surgeons that have committed higher resource utilizations.

However, as we have discussed in Chapter 2, such implicit performance indicator as "utilization rate" inherently excludes the "efficiency" consideration from the capacity planning. Specifically, it is rather possible that departments or surgeons who work extremely inefficiently and/or whose surgeries incur large fiscal loss to the hospital keep receiving high levels of OT capacity in every revision simply because they occupy the OT for the longest durations. After discussions with the hospital, it is determined that the "profit" and the "patient throughput" will be used as performance metrics in the new OT capacity planning. Recall our discussion in Chapter 1, these two objectives are set in line with the two (inherently conflicting) "fiscal" and "social" goals of the hospital. We consider the setting up of the objectives according to the hospital's ultimate goals as the first step towards rationalization of OT capacity planning.

Furthermore, it should be noted that by setting up two objectives in the planning, instead of a single optimal solution, the hospital will be provided with a spectrum of OT capacity plans (i.e. each corresponding to a better performance than other plans in one objective, and a worse performance in the other). Flexibility will thus be allowed for the hospital decision maker to balance out the trade-off between their two objectives and pick out a most satisfactory design from all candidate plans.

From our discussion with the hospital, it is also decided that our new OT capacity planning will be conducted in a “centralized” manner, i.e. allocation of slots will be done directly to the surgeons within each clinical department, as opposed to the current two-stage practice in the hospital. As one could foresee, this direct allocation may encounter administrative difficulties if new doctors arrive and existing doctors leave the hospital frequently (i.e. a lot of modifications on the OT plan will have to be carried out regularly in such cases). However, by allocating the capacity in the centralized manner, transparency and the perception of fairness among all the surgeons in the hospital can be achieved. Also, the clinical departments can be relieved from the work of inner-department capacity allocations and the possible hassles associated with these, to focus fully on their medical performances. We thus consider our centralized capacity allocation generally a right move.

Additionally, in the previous OT practice, there exist certain departments which do not have the luxury to provide every one of their surgeons with a dedicated OT slot. In such departments, surgeons who are not allocated any OT slots (usually those with relatively small patient loads) will have to “borrow” capacity from the other surgeons (i.e. seek for approval and book their cases under the other surgeons’ slots) or wait until the booking cut-off time (3-4 days before surgery date) to book their surgeries. It is perceived rather unfair by these surgeons that all the slots are assigned to only those “major” surgeons, and frustrations have been incurred for these “disadvantaged” surgeons at work.

In order to resolve such hassles, in our new capacity planning we set up a securing "shared" slot between the surgeons in each department. Specifically, we allow to leave one OT slot to be "shared" in each department among the surgeons who are not provided with any dedicated slot(s). Once a surgeon is allocated dedicated slot(s), s/he would not quality for using this shared slot. Depending on the number of surgeons and number of total allocated slots, this shared slot may or may not be applicable for every department. We anticipate that the implementation of this shared slot will reduce the hassles among the surgeons in fighting for the OT capacity in the future practice.

In deriving the new capacity allocation, there are several practical concerns/ constraints that have to be respected according to the hospital. First of all, it is required that each clinical department maintains sufficient capacity to serve a "minimum number of patients per week". This constraint is set to keep the hospital's comprehensive capability to conduct surgeries for all types of patients, and is maintained regardless of the profitability of each clinical department. On the other hand, an "upper limit" is required to be set on the number of slots a surgeon could receive. The purpose of this is to maintain an overall perception of fairness among the surgeons. Lastly, it is ruled that all the hospitals' surgeons must be arranged in the new capacity allocation scheme, i.e. any surgeon must be arranged into at least one slot, either "dedicated", or "shared".

Chapter 4: Simulation Module

After investigating the current capacity allocation practice at the hospital's OT and proposing a new approach to it, we attempt to develop an algorithm that derives the optimal OT capacity allocations for the hospital. As we have discussed, the revised goals of the capacity allocation approach is to achieve the maximized "profit" and "patient throughput" for the hospital. In addition, contrary to the old practice where every OT slot is given to only one surgeon, in the new system we would allow each department to have one OT slot that is shared among all its surgeons who are not allocated dedicated slot(s).

Prior to developing the algorithm, we are required to first identify an "evaluation" method for our OT solutions. Specifically, certain approach is needed that can translate for us any OT plan (as characterized by the full set of allocation of all the OT slots to all the surgeons) into the performance measures we are interested in, i.e. the "number of patients" served, and the "profit" generated. In addition, in line with the nature of our problem where the OT capacity is planned for a week, in evaluating the OT plans it is desirable that we express the performance of the plans in the same manner, i.e. as the weekly patient throughput, and the weekly profit generated.

In order to accurately predict the patient throughput and the profit of any given OT plan, it is required that we understand and model in details the complex dynamics how the patients' surgery requests are received, handled, and ultimately served (or cancelled) in the OT system. Because of the obvious advantage of computer simulation in modeling the dynamics of complex systems and in incorporating empirical data, we develop a simulation model to mimic the OT operations and to evaluate the OT plans.

In this chapter we will first explain how we decide on the scale of our simulation model. This is followed by a discussion on how we prepare the input data for the model. After these, we will respectively present the structure, the dynamics, and the assumptions of our simulation model.

4.1 Scale of Simulation Model:

In designing the simulation model, we have noticed the fact that any "full" OT plan (i.e. allocation of all the OT slots to all the surgeons in the hospital) can essentially be decomposed into a set of "elements", each being characterized by the allocation of "certain number of slots" to "certain number of surgeons" (See Figure 4.1).

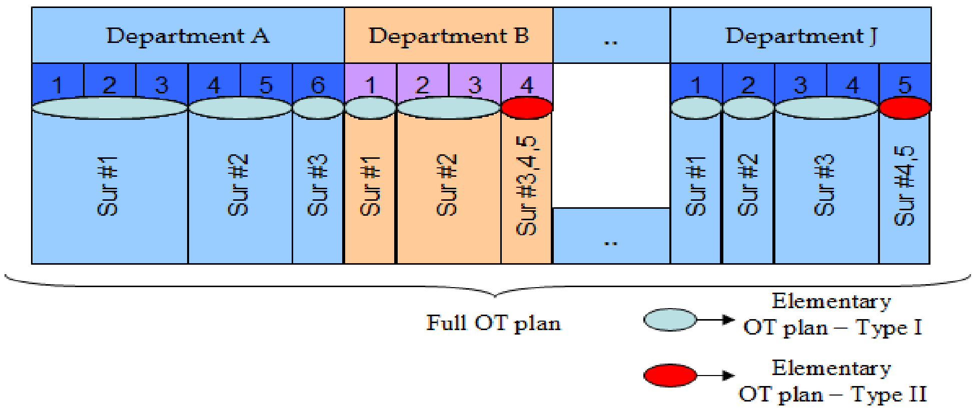

Figure 4.1: Decomposition of a Full OT Plan into its Elements

More specifically, there are two possibilities of the "elementary" OT plans. One of these is the case where several OT slots are dedicated to a single surgeon. In such cases, the number of slots allocated will be greater or equal to one, but it has to be below the upper limit on the maximal personally capacity as set by the hospital. The other type of elementary OT plan is the case where several surgeons together share one slot. As we have stated, in each department there can be one and at most one shared capacity slot.

Figure 4.1 illustrates the idea of composing a full OT plan into its elements. Note in this figure the presence of the two types of elements in a full OT plan.

Bearing in mind that our new approach to allocating the OT capacity has guaranteed each surgeon with at least one OT slot (either dedicated or shared), thus no "borrowing" or "sharing" of slots should take place among any surgeons in the future operations. Because of this, it can be seen that all the "elements" in a full OT plan are essentially "mutually independent" to each other in terms of their contribution to the performance of the full OT plan.

Therefore, in our study we decide to focus on developing a simulation model that evaluates the performance of an "elementary" OT plan. Being more specific, we aim to develop such a "generic" simulation model that can mimic the procedure how, when "any combination of surgeon(s)" are allocated "any number of OT slot(s)", the surgery requests for these surgeon(s) appear, get booked, and subsequently get conducted or cancelled in the OT system. After running this model, the performance of any elementary OT plan can be evaluated in terms of its weekly "profit", and "patient throughput" generated for the hospital. The performance of any full OT plan will thus be evaluated through the procedure of (1) decomposing the full plan into all its elements, (2) evaluating each of these elements individually, and (3) integrating (i.e. summing up) the obtained performances of all the elementary plans.

4.2 Data Preparation:

After understanding the requirements for our simulation model, we proceed to preparing the input data. We identified five categories of data that are essential to characterize the various aspects of the surgeries' booking and execution process. Specifically, these five datasets include the "arrival time" of the patients, the "duration", and the "profit" of the surgeries conducted by the different surgeons as well as the "cancellation" pattern and the "queue-abandonment" behavior of patients under each clinical department.

The arrival time data generally refers to the "arrival rate", and the "inter-arrival pattern" of the patients.

The arrival rate of the patients is estimated based on the surgeons' "caseloads" in the period of study (a 9-month duration). This information is available from the hospital's IT system. In extracting the caseload for the individual surgeons, we have paid particular attention to the difference between the "realized", and the "actual" patient demand of each surgeon in the study period. In particular, in line with our concern on the patients' "queue-abandonment" behavior, we were also aware of the fact that each surgeon more or less had lost certain patients due to their intolerance of the long waiting time, and thus the caseload we observe on him/her in the study period is only a portion of his/her "real" patient demand. Therefore, we have consulted the administrative staff who handle the surgery booking in each clinical department for a rough percentage of “abandoned patients” over the long-run. Based on this information, we have estimated for each surgeon in that department (by assuming that all surgeons in the same department had the same percentage of lost patients) his “real caseload” in the study period.

Due to the lack of information, we did not have the luxury to investigate the patterns on the patients’ “inter-arrival times”. However, as has been reported in many other settings in the service industry, customers who arrive “completely randomly” into a system depicts such a “memory-less” property on their inter-arrival time, which is held within the Poisson arrival pattern. We have thus made an assumption that the patients in our hospital arrive also with the Poisson pattern into the OT system.

The “durations” and the “profits” of all the elective surgeries conducted by all the surgeons in the study period are also available in the hospital’s IT system. We have thus built a data file for each surgeon in which detailed information concerning all his “historical” surgeries is stored. Specifically, this information covers the “Patient’s Class” (i.e. private or subsidized), the “Surgical Code” (i.e. nature of the surgery), the “Duration”, and the “Profit” for each case of this surgeon in the study period.

Furthermore, in creating the data file for each surgeon, the “duration” and the “profit” information stored in each surgery entry are calculated as the average values of all those same cases (i.e. cases with the same “surgical code”) as performed by this surgeon. This step is done because of our awareness that our simulation model needs to mimic the “actual” system characteristics in order to be useful. Specifically, in practice when a surgeon books a case s/he would not know for sure how long it will actually take. The best estimation to the surgery time would therefore be the average duration of all the similar cases s/he had done before. As for the profit, it is more computational- efficient to directly assign the average values to each type of surgery. This is because when the simulation is run long enough, and when each type of surgery (say with different profits initially assigned) appears for a large number of times, the average profit of these simulated cases is going to converge just to the average profit of these cases if each is assigned at the very beginning. It thus saves the time for the simulation to converge by assigning average profits to the cases.

Information concerning the surgery cancellations is also obtained through surveying the administrative staff from each department. In particular, the “cancellation pattern” of patients under each department is characterized by three factors, namely, the total “percentage” of cancellation; the “distribution” of the cancellations throughout the waiting time (in other words, the “odds” a case is cancelled in the each week before the scheduled surgery date); and lastly, the percentage of cancelled cases that will “revisit” the hospital for operation.

Lastly, with respect to the patients’ “queue-abandonment” behavior, we have again consulted people in each department for their rough idea “how much of waiting their patients could tolerate without abandoning for another hospital”. The answers are taken in “distributions”, i.e. in the format of tables recording the informed “percentage of abandonment” against each level of “waiting time”. It needs to be noted that in here the waiting times are measured in “weeks”, because of the fact that the surgeries are all booked into the “earliest possible week” in the surgeon’s slot(s). In addition, as suggested by the staff, a differentiation is further made between the private and the subsidized patients in each department on their "waiting tolerance level".

4.3 Simulation Model Structure and Dynamics:

Our simulation model consists of two major components, namely, the "surgery generation module", and the "surgery booking/execution process module".

The general structure of the simulation model is as follows:

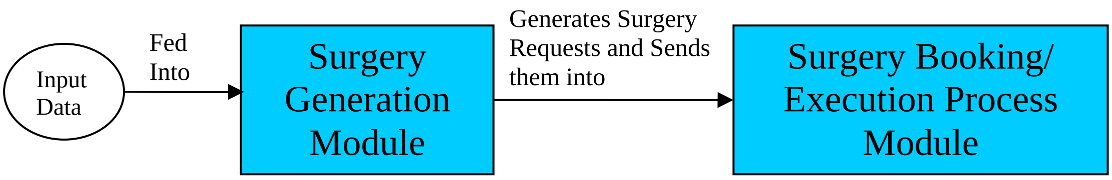

Figure 4.2: General Simulation Model Structure

The "Surgery Generation Module" is the module that imports the input data files into the model at the beginning of simulation; and afterwards, it controls the generation of surgery requests according to information stored in the data. The requests generated will be sent to the downstream "Surgery Booking/ Execution Process Module" to get served. In addition, the module also controls the execution/termination of each simulation replication, as well as summarizing and printing simulation output.

The "Surgery Booking/ Execution Process Module" is the module that mimics the exact procedure how, upon its appearance, a surgery request gets booked, handled, and ultimately executed or cancelled in the OT system.

When simulating each scenario (i.e. each elementary OT plan), the rate how fast the "Surgery Request Generation Module" creates the surgery requests is decided by two factors: (1) the total patient load of all the surgeon(s) involved in the plan in the period of study, and (2) the estimated "lost patient load" of all the surgeon(s) in the corresponding period. The statistical distribution used to model the inter-arrival times is exponential.

Upon generation of a surgery request, information concerning the case's "Surgeon", "Department", "Patient Class", "Duration", and "Profit" will be assigned according to one of the historical cases of the surgeon(s) that is picked randomly. In addition, knowing the "department" and "patient class" of the case, the "cancellation pattern", and the "queue-abandonment feature" (different between "subsidized" and "private" patients) will also be assigned to the case.

The generated surgery request will then be passed down to the "Surgery Booking/ Execution Process Module". The following procedure how the case gets booked, handled, and eventually executed in the simulation model is summarized in Figure 4.3.

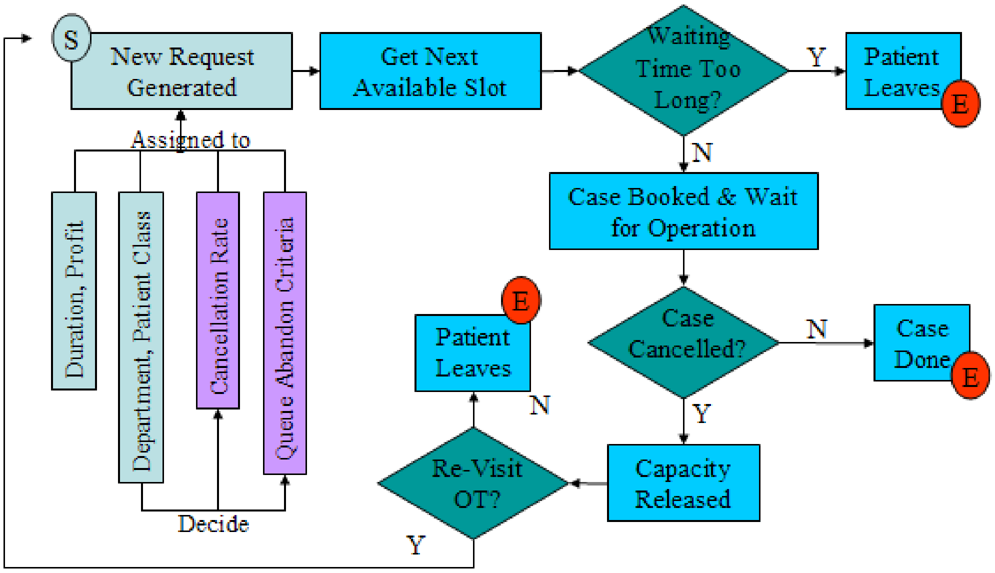

Figure 4.3: Surgery Booking/Execution Process Dynamics

Specifically,

- Based on the estimated surgery duration, the next available week that has sufficient time for the case will be found, and the case will be scheduled then. However, if the waiting time is too long, i.e. longer than stated in the "abandonment" attribute, then the request will leave the system directly.
  - In here, due to the fact that our capacity planning does not specify the exact time in a week an elementary OT plan will be arranged, in modeling the patients' waiting time we have adapted a more "conservative" measure.
  - To be specific, we have set the rule that any surgery that cannot be operated within the current week will be considered to have rendered a "whole-week" waiting time for the patient. The same extension on the waiting time is also applied to cases that cannot be operated within two weeks, three weeks, and so forth.

    - Subsequently, it is this conservatively estimated waiting time that is compared to the patient's queue-abandonment attribute to decide if he/she would leave the system. By doing so, we are more confident that our simulation model would not over-estimate the performances, i.e. the patient throughput, and largely depending on this, the profit, of the OT plans.
- Throughout the waiting, the case could be cancelled according to the "cancellation pattern" assigned to it upon its arrival.
    - More specifically, after each week's waiting the "cancellation probability" in that week (characterized by the number of extra weeks before surgery) will be looked up, and a decision is randomly made (through coin tossing according to the cancellation probability) whether the case is to leave the system or to stay.
- When a case is cancelled, the corresponding capacity that is previously reserved for the case will be freed-up, and will be open for booking the following newly-arrived patients. In addition, if the cancelled case will re-visit the OT, it will be treated as a new surgery request.
- When a case is successfully executed, before it leaves the system, the "patient throughput" and the "profit" will be updated for the scenario under simulation.

At the end of each simulation replication, the desired simulation outputs, namely, the "weekly profit", and the "weekly number of patients served", will be exported by the "Surgery Request Generation Module".

4.4 Simulation Model Assumptions:

A few assumptions are made in developing the simulation model. These are summarized as follows:

1. The arrival process of the surgery requests is assumed to be Poisson, and the arrival rate is assumed to be consistent over time (rendering a homogeneous Poisson process).
2. Surgeries are booked “side-by-side” in the capacity slots (i.e. there exists no “holes” in between), and therefore the allocated capacity is deployed “optimally”. However, note that this assumption is not to say that every minute in an OT block will be used up to book surgeries. Only surgeries that fit into the remaining block time shall be booked; otherwise the planner will automatically seek the next weeks’ blocks, leaving an open space in the current week. Such open spaces, nevertheless, can still be utilized for booking latter-arrived surgery requests if they indeed fit into it.
3. When the slot(s) are shared by multiple surgeons, no priority is given to any surgeon in booking their cases. All the surgery requests are served on the First-Come-First-Served basis.
4. All patients under the same department are assumed to have the same “Cancellation Pattern” and “Queue Abandonment Behavior”.
5. For the simplicity in implementing the simulation model, it is assumed that each patient has a “deterministic” tolerance level for his surgery waiting time.
6. Similarly, in modeling the surgery cancellations, it is assumed for all the clinical departments that the probability of a case’s cancellation is uniformly distributed

over all the weeks of the case's waiting time. This simplification could greatly reduce the number of parameters to feed into the simulation.

Chapter 5: Model and Optimization Algorithm

Based on the simulation model that evaluates the OT plans, we propose an algorithm to identify the optimal OT capacity allocation for the hospital. In this chapter, we will first present a formal formulation of our capacity planning problem. After this, we explain the reasoning and the mechanism of our algorithm in details. The implementation and numerical results of the algorithm will be presented in the next chapter.

5.1 Problem Formulation:

We have formulated the OT capacity planning problem at the hospital as follows:

Indices and sets

i index for departments; $i \in D$

j index for OT plans under each department; $j \in P_i$

k index for surgeons under each department; $k \in S_i$

D set of departments to be allocated OT capacity

P_i set of departmental OT plans considered for department i

S_i set of surgeons in department i

Parameters and data

r_{ij} profit generated as assigning plan j to department i

t_{ij} patient throughput generated as assigning plan j to department i

a_{ijk} number of dedicated slots the surgeon k is allocated in plan j in department i

$$b_{ij} = \begin{cases} 1 & \text{if plan j in department i consists of a shared slot} \\ 0 & \text{if plan j in department i consists all of dedicated slot(s)} \end{cases}$$

CT total number of capacity slots to be allocated

$MaxCap_i$ upper limit on number of slots allocated to department i

$MinPT_i$ lower limit on patient throughput for department i

MaxCap_P upper limit on number of slots allocated to any single surgeon

Decision Variables

$$x_{ij} = \begin{cases} 1 & \text{if plan j is assigned to department i} \\ 0 & \text{otherwise} \end{cases}$$

Objective

Maximize (z_1, z_2), where:

$$z_1 = \sum_{i \in D} \sum_{j \in Pi} r_{ij} * x_{ij} \text{ (total profit)}$$

$$z_2 = \sum_{i \in D} \sum_{j \in Pi} t_{ij} * x_{ij} \text{ (total patient throughput)}$$

Constraints

$$\sum_{j \in Pi} (b_{ij} + \sum_{k \in Si} a_{ijk}) * x_{ij} \leq MaxCap_i \qquad \text{for all } i \in D \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$$\sum_{j \in Pi} t_{ij} * x_{ij} \geq MinPT_i \qquad \text{for all } i \in D \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

$$\sum_{j \in Pi} a_{ijk} * x_{ij} \leq MaxCap\_P \qquad \text{for all } i \in D \text{ and all } k \in S_i \ldots\ldots\ldots(3)$$

$$\sum_{i \in D} \sum_{j \in Pi} (b_{ij} + \sum_{k \in Si} a_{ijk}) * x_{ij} \leq CT \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

$$\sum_{j \in Pi} x_{ij} = 1 \qquad \text{for all } i \in D \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5)$$

In our problem formulation, the building blocks of the overall capacity allocation are the capacity plans on the “departmental” level. Specifically, for each clinical department we maintain a set of “departmental OT plans”, each of which is characterized by a “capacity level” (i.e. total number of slots allocated for this department) as well as a detailed allocation of these slots to all the surgeons in the department.

The basic task of the model is therefore to pick from among many of such departmental OT plans one plan for each department and to form the optimal “global” (i.e. “hospital-wide”) OT plans.

The method how the departmental OT plans are generated lies in the algorithm portion, which will soon be explained in the next section. But note that once a departmental OT plan is specifically known, three things can be known at the same time. These include: (1) the number of capacity slots each surgeon in the department is granted, (2) the fact whether this plan consists of a shared slot or is made up purely by dedicated slots, and (3) the profit and the patient throughput the plan will generate (through simulation as will be discussed soon). These three sets of information form the parameters for the model.

The other parameters for the model include the total number of slots being allocated, the maximum capacity allowed, and the minimum patient throughput that must be maintained in each department as well as the maximum capacity allowed for any single surgeon in the hospital.

The two objective functions in our problem formulation respectively seek to optimize the "total patient throughput" and the "total profit" for the hospital. These statistics are derived by summing up the parameters associated with all the departmental OT plans picked for each and every one of the departments. Among the five constraints listed, the first and the second constraints respectively control the "maximum capacity" each clinical department is entitled to receive and the "minimum patient throughput" it must maintain. The third constraint limits the maximum capacity for each individual surgeon in any department. The fourth constraint controls the total amount of OT capacity being allocated (i.e. the physical capacity limit), and the last one ensures that only one design is picked in each clinical department.

5.2 Optimization Algorithm:

In line with the formulation of our problem where the overall OT capacity allocation is decomposed into the various "departmental" OT plans, we develop a two-phase algorithm for solving the model, the two phases of which respectively seeks to identify the optimal "departmental", and the optimal "hospital-wide" OT capacity plans.

More specifically, the Phase One of our algorithm aims to prepare, for each clinical department, one set (i.e. the set P_i) of "feasible" and "local-optimal" OT plans, such that all the constraints set on the departmental level are met and that the maximal profit and maximal patient throughput are generated for the department. Upon obtaining such sets of candidate OT plans for each department, the Phase Two of our algorithm shall pick from among these candidates one plan for each department and form the optimal set of “hospital-wide” OT capacity plans.

Due to the fact that we have two objectives under consideration, our OT plans, both on the departmental and on the hospital level, are expressed by Pareto sets.

We develop a greedy-based searching algorithm for the Phase One which consists of 8 major steps; and for Phase Two, we build an Integer Programming (IP) model to pick the best combinations of the departmental OT plans. We shall now describe our searching algorithm and the IP model in details.

5.2.1 Algorithm Phase One – Generation of Departmental OT Plans:

We develop a breadth-first greedy search algorithm to prepare the candidate OT plans for each department. Specifically, our algorithm starts from the “one-slot scenario” for each department, and it later proceeds to the higher capacity levels through examining the effectiveness of all the possible extensions on its current solutions.

On each capacity level, only the “best” solutions are kept in order to reduce the searching space in the next iterations. As mentioned before, the best solutions are expressed as a Pareto set, which consists only of solutions that are non-dominated by any other solution in the set in terms of both performance criteria (i.e. profit and patient throughput). In addition, in trying to create new solutions on each capacity level, the constraint set on the “maximum dedicated capacity” for each surgeon is directly incorporated, i.e. no solution is generated that violates this constraint.

The algorithm stops when no further improvement is made on either of the two objectives as a department is offered one more slot, or when the department's total capacity reaches the upper limit set for it.

After the algorithm stops, the solutions obtained in all the previous capacity levels that violate the "minimum patient throughput" constraint for the department are deleted. The remaining solutions are subsequently exported as the set of candidate departmental OT plans.

Detailed steps of the searching algorithm are as follows:

Input: S_i, $MaxCap_i$, $MinPT_i$, MaxCap_P

Output: P_i

Step 1. (Initialization) Set capacity level $\alpha = 1$, and set $P_i = \emptyset$.

Step 2. (Build and Evaluate Starting Plan) Set the plan set at capacity level 1, P_i^1, to the single plan whereas "the entire surgeon set S_i shares the single capacity slot". Simulate to obtain the objective values r_{i1}^1 and t_{i1}^1 of this plan.

Step 3. (Generate New Feasible Plans) Set $\alpha = \alpha + 1$;

Based on every plan $j^* \in P_i^{\alpha-1}$, generate one new solution for set P_i^{α} at a time by allocating one additional slot to:

- Any surgeon $k \in S_i$ who is in the "shared slot" in j^*. By doing this, the surgeon k will be out of the shared slot as it appears in the new plan j,
- Any surgeon $k \in S_i$ that has dedicated slot(s) in j^*, yet $a_{ij^*k}^{\alpha-1} \leq$ MaxCap_P-1;

Afterwards, delete the redundant identical plans in P_i^{α} (these are generated from different $j^*s \in P_i^{\alpha-1}$).

Step 4. (Evaluate New Plans) For every $j \in P_i^{\alpha}$, simulate to obtain the objective values r_{ij}^{α} and t_{ij}^{α}.

Step 5. (Build Pareto Set) For every two plans j and $j^* \in P_i^{\alpha}$, if $r_{ij}^{\alpha} \leq r_{ij*}^{\alpha}$ and $t_{ij}^{\alpha} \leq t_{ij*}^{\alpha}$, delete plan j from set P_i^{α}.

Step 6. (Stopping Criteria) If α =MaxCap$_i$, then go to **Step 7**;

If every plan $j \in P_i^{\alpha}$ is dominated by some plan $j^* \in P_i^{\alpha-1}$, then go to **Step 7**;

Otherwise, go to **Step 3**.

Step 7. (Build Set P_i) For every $\alpha \in$ [1..MaxCap$_i$] and every plan $j \in P_i^{\alpha}$, if $t_{ij}^{\alpha} \geq$ MinPT$_i$, put j into the output plan set P_i.

Step 8. Export P_i.

The major mechanism of the searching algorithm (Steps 1 through 6 is summarized in the Figure 5.1 below.

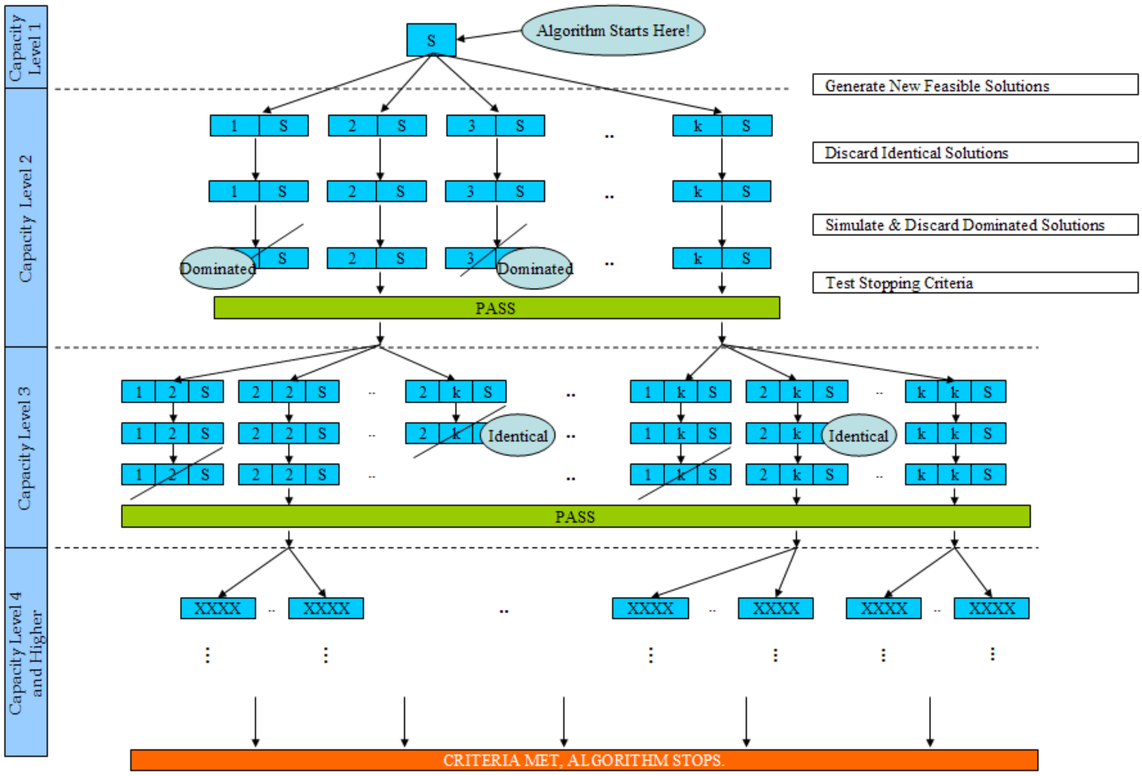

Figure 5.1: Major Mechanism in Phase One Searching Algorithm

5.2.2 Algorithm Phase Two – IP Model:

Note that the Phase One of our algorithm has tackled the majority of constraints involved in our capacity planning problem. These include the maximum capacity and the minimum patient throughput for each department, and the maximum capacity for each individual surgeon in the hospital. The IP model that picks the best combinations of the departmental OT plans is simply formed as follows:

Decision Variables

x_{ij} (binary) 1 if department i is assigned to plan j, 0 otherwise

Objective

Maximize (z_1, z_2), whereas:

$$z_1 = \sum_{i \in D} \sum_{j \in Pi} r_{ij} * x_{ij},$$

$$z_2 = \sum_{i \in D} \sum_{j \in Pi} t_{ij} * x_{ij}$$

Constraints

$$\sum_{i \in D} \sum_{j \in Pi} (b_{ij} + \sum_{k \in Si} a_{ijk}) * x_{ij} \leq CT$$

$$\sum_{j \in Pi} x_{ij} = 1 \qquad \text{for all } i \in D$$

To solve this bi-objective IP, we develop an approach that limits and gradually relaxes one of the objective functions. Specifically, we introduce an additional constraint on the "total patient throughput" of the global OT plan:

$$\sum_{i=1}^{D} \sum_{j \in Pi} x_{ij} * t_{ij} >= TP \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

We first set TP (i.e. our arbitrary constant) to a rather low level, and run the IP model to identify the optimal combination of the departmental plans that maximizes the profit. Afterwards, we gradually increase the value of TP whilst examining the impact of this on the model's choices of departmental plans as well as on the value of the other total profit objective.

Chapter 6: Algorithm Implementation and Numerical Results

In this chapter, we explain how we implement the two-phase algorithm and show the numerical results as we apply the algorithm to solve our particular OT capacity planning problem. We first summarize the parameters describing our system being solved; and after this, we describe for each of the two phases of our algorithm the implementation methodology and the numerical results.

6.1 System Parameters:

Our OT consists of 10 elective ORs, and thus a total of 100 OT slots (i.e. CT=100) to be allocated in a week. Altogether, there are 10 clinical departments (i.e. D=10) that require the OT capacity.

The Table 6.1 summarizes the "number of surgeons" (i.e. the S_i's), and the parameters for both the "maximum capacity" (i.e. the $MaxCap_i$'s) and the "minimum patient throughput" (i.e. the $MinPT_i$'s) constraints set for each clinical department. These constraint parameters are obtained through consulting the hospital, and they are decided mostly based on the original capacity plan at the OT, combined with the overall OT utilization rates and the total patient loads of the departments in the studied period. The ordering of the departments is based on the number of surgeons in them.

Table 6.1: System Parameters for the Departments

Department (i)	Number of Surgeons (S_i)	Max Capacity ($MaxCap_i$)	Min Patient Throughput ($MinPT_i$)
Department #1	**18**	**32**	**24**
Department #2	**13**	**34**	**26**
Department #3	10	11	13
Department #4	6	11	10
Department #5	5	20	11
Department #6	4	10	8
Department #7	3	8	8
Department #8	1	4	3
Department #9	**1**	**1**	**N.A.**
Department #10	**1**	**1**	**N.A.**

In the table, it is revealed that the top two departments in terms of their surgeons' numbers also correspond to rather large parameters on both the "minimum patient throughput" and the "maximum capacity" constraints as compared to the other departments. In particular, it can be derived that these two departments are expected to generate almost half of the hospital's patient throughput, and are expected to occupy almost half of the hospital's OT slots. These data are indicative of the extremely important roles played by these two departments in the OT system.

On the other hand, Departments #9 and 10 have only one surgeon each, and they both have appeared to use the OT rather rarely in the studied period. After discussing with the hospital, it is decided that one capacity slot will be allocated to each of these two departments.

Moreover, the maximum capacity constraint set for all the individual surgeons in the hospital is decided to be 4 slots (i.e. MaxCap_P=4).

Because of the elimination of Departments #9 and 10 from consideration, in applying our algorithm we shall work only on allocating the rest 98 slots to surgeons from the rest 8 departments, i.e. Departments #1-8.

6.2 Phase One Algorithm Implementation and Results:

The searching algorithm in Phase One is directly embedded into our simulation model. Specifically, an additional “algorithm” module is created on top of the existing model which implements all the logics in the algorithm, and governs the execution of the program for each clinical department.

Note that any departmental OT plan is essentially also a combination of “elementary” OT plans each allocating certain number of slots to certain surgeons in the department (refer to Figure 5.1), in order to enhance the efficiency in evaluating the departmental OT solutions, a “simulation result table” is retained in the algorithm. In the table, the performances of all the elementary plans that have been evaluated are stored. Therefore, through the algorithm’s execution, whenever a new departmental OT plan is generated, the program would firstly look up the table to see if any part(s) of this new plan has been evaluated. Simulations runs are only conducted to the new elementary plans.

In addition, in running the simulations we designate a “warm-up” period of 1,000 simulation weeks for each scenario (i.e. each elementary OT plan). After finishing the warm-ups, each system is run for additional 10,000 weeks to get evaluated. Five replications are conducted to each simulated scenario in order to rule out the biasness on the estimated system performances rendered by randomness.

In this section, we take a relatively small department, namely the Department #6, as an example to demonstrate the execution process and the numerical results of our searching algorithm.

The Department #6 has 4 surgeons, whose “Total Profit”, “Total Caseload”, and “Total Surgery Duration” (all data normalized) in the study period are summarized in Table 6.2.

Table 6.2: Study Period Information for Department #6 Surgeons

Surgeon	Total Profit	Total Caseload	Total Duration
1	$111,766	133	203
2	$93,922	166	207
3	$44,310	78	105
4	$86,113	80	127

Table 6.2 reveals that among the four surgeons in the department, surgeons #1, 2 and 4 had brought significantly higher levels of profit for the hospital as compared to surgeon #3. In addition, among these three most “profitable” surgeons, surgeon #4 had achieved his profit by consuming only 127 time units of the OT capacity. This statistic corresponds to almost only half of the rest two surgeons’, and is strongly indicative of surgeon #4’s “efficiency” in generating the profit. Surgeon #2 had the largest caseload among the four surgeons; however, he had also consumed the largest amount of time in the OT. Surgeon #1 had the largest profit, second largest caseload, and had consumed the second largest amount of OT time among the four surgeons. Surgeon #3 does not seem to have any “unique” character in any aspects.

The Phase One algorithm is run for Department #6, and the Pareto solutions generated on all the capacity levels are summarized in Table 6.3, together with their performances in terms of both of the two objectives.

Table 6.3: Pareto Solutions with Performances at all Capacity Levels - Department #6

Cap Level	Pareto Solutions							
	Cap Allocation				Weekly Profit		Weekly Pat. TPT	
	Sur #1	Sur #2	Sur #3	Sur #4	Mean	Stdev%	Mean	Stdev%
1	.	.	.	.	$1, 631	1. 04%	2. 21	0. 68%
2	1	.	.	.	$2, 627	0. 88%	4. 26	0. 59%
	.	1	.	.	$2, 747	2. 04%	4. 09	1. 02%
	.	.	.	1	$2, 901	0. 70%	3. 79	0. 66%
3	1	1	.	.	$3, 744	1. 11%	5. 99	0. 56%
	.	1	.	1	$3, 850	0. 67%	5. 80	0. 72%
	1	.	.	1	$3, 961	0. 92%	5. 71	0. 58%
	2	.	.	.	$3, 998	1. 24%	5. 31	1. 10%
4	1	2			$4, 653	1. 09%	7. 58	0. 77%
	1	1	1	1	$4, 773	0. 90%	7. 39	0. 79%
	2	1	.	.	$4, 996	1. 02%	7. 20	0. 58%
	2	.	.	1	$5, 212	0. 75%	6. 92	0. 24%
5	1	2	1	.	$5, 682	1. 09%	8. 98	0. 93%
	2	2	.	.	$5, 905	1. 00%	8. 79	0. 66%
	2	1	1	1	$6, 024	0. 64%	8. 60	0. 78%
	2	.	.	2	$6, 236	1. 23%	7. 66	0. 33%
6	2	2	1	1	$6, 934	0. 86%	10. 19	0. 82%
	2	1	1	2	$7, 048	1. 11%	9. 34	0. 62%
7	2	3	1	1	$7, 625	1. 14%	11. 25	0. 59%
	2	2	1	2	$7, 957	0. 99%	10. 93	0. 69%
8	2	3	1	2	$8, 649	1. 26%	11. 99	0. 63%
	3	2	1	2	$8, 681	1. 01%	11. 62	0. 57%
9	3	3	1	2	$9, 372	1. 27%	12. 68	0. 79%
10	3	3	2	2	$9, 592	1. 09%	13. 14	0. 70%

In the table, the number of OT slots allocated to each surgeon in each Pareto solution is stored in the column under that particular surgeon. Additionally, in representing the shared slots, we mark those surgeons who are arranged in such a slot with dots (“.”). For example, the four dots in the only solution on capacity level one imply that all the four surgeons are put into the shared slot. And in the first Pareto solution on capacity level 2, the three dots in the columns of surgeons #2, 3, 4 imply that these are the surgeons who have to share one capacity slot.

The performances of the OT plans are evaluated by the average “weekly profit” and the average “weekly patient throughput” of all simulations runs. The precision levels of the simulation results, as measured by the “percentage of the *standard deviations* over the *mean* of the objective values”, are mostly within 1%. This gives us the confidence to say that throughout executing the algorithm, the probability that certain “non-dominated” solutions were wrongly deleted is rather small.

In Table 6.3, it is shown that as the algorithm proceeds from the first capacity level (i.e. the scenario where all surgeons share a single slot) to the “two-slot” capacity level, the surgeons #1, 2, 4 are each picked out of the shared slot to form 3 new plans. Moreover, throughout the higher capacity levels where the algorithm progressively explores the possibilities to allocate additional capacities, the Pareto solutions rendered are mostly characterized by adding dedicated slots for surgeons #1, 2, and 4. As we have examined, the three surgeons #1, 2, and 4 are the most “profitable” surgeons in the department. Furthermore, Surgeons #1 and 2 consist of the largest caseloads, which convert to the largest contributions of patient throughputs for the department. It is also depicted in the table that these two surgeons in general are allocated more dedicated slots as compared to surgeon #4 s over the various capacity levels.

In summary, the running results of the algorithm for Department #6 have suggested that our searching algorithm is able to successfully identify the "important surgeons", pick them out of the "shared group" and arrange them into the dedicated capacity slots.

The performances of the Pareto solutions generated by the algorithm on the various capacity levels are further plotted in Figure 6.1.

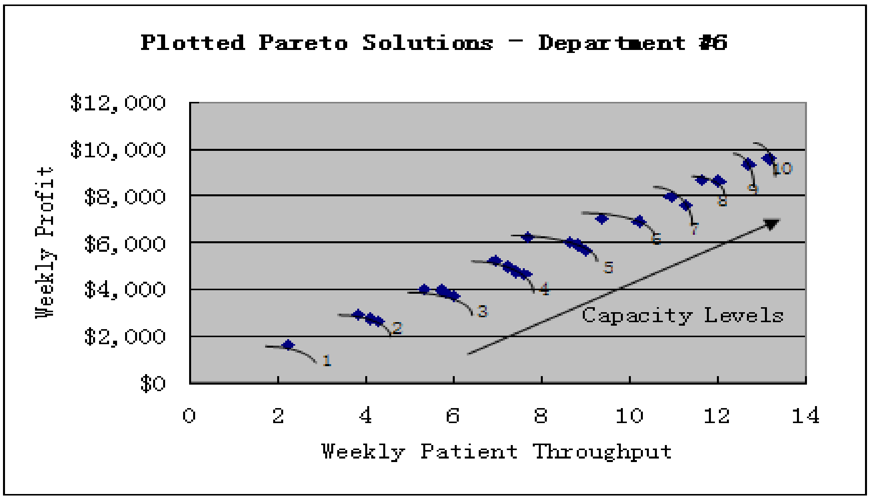

Figure 6.1: Plotted Pareto Solutions - Department #6

A few observations are made on Figure 6.1. First, it can be seen that the plotted performances of the Pareto solutions on any capacity level would always fall "outside" of those solutions on the previous capacity level(s). The reason for this is that when the algorithm proceeds between capacity levels, any solution generated based on a Pareto solution on the lower capacity level would dominate this "parent" solution (note that no difference exists between a solution and its “parent” except that one surgeon is given an additional dedicated slot); however, this dominated “parent” solution is non-dominated by any of the other Pareto solutions on the lower capacity level. It can thus be implied that no solution in the Pareto set of a higher capacity level could possibly be dominated by any solution in the Pareto set of a lower capacity level.

Moreover, it is also revealed in Figure 6.1 that the “gaps” between the performances of the Pareto sets generated on the different capacity levels have the tendency to diminish, and eventually to eclipse as the capacity level goes up. This is due to the fact that the marginal “profit” and/or the marginal “patient throughput” of any surgeon in any department would decrease as s/he is offered more capacity slots. Intuitively, when all the surgeons in a department are provided with sufficient capacity to handle all their cases, it would not make any different if any additional slot is given to this whole department.

At last, after truncating the infeasible designs with respect to the minimum patient throughput constraint, the algorithm has identified a total of 13 “Candidate OT plans” for Department #6.

The algorithm is run for all the eight clinical departments, and the patterns on the plotting of each department’s Pareto solution performances have appeared to be similar. Figure 6.2 presents the plotted Pareto solution performances in Department #1, the department with the largest number (18) of surgeons.

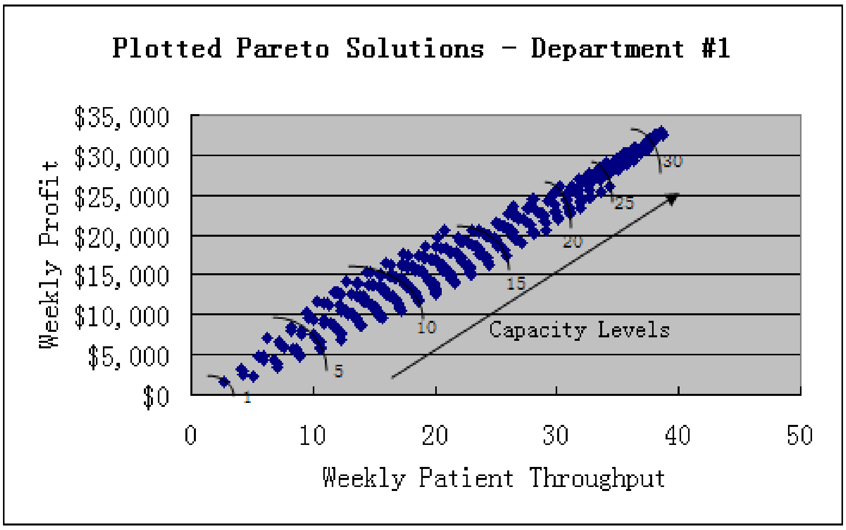

Figure 6.2: Plotted Pareto Solutions- Department #1

The following Table 6.4 summarizes the number of candidate OT plans obtained for each department at the end of Phase One. It can be seen that the solution space of a department depends largely on its “scale”, whereby Departments #1 and 2 have comprised over 75% of the total number of candidate plans

Table 6.4: Summary of Phase One Results – Number of Candidate OT Plans

Department (i)	Number of Surgeons (S_i)	Max Capacity ($MaxCap_i$)	Min Patient Throughput ($MinPT_i$)	Number of Candidate OT Plans
Department #1	18	32	24	**165**
Department #2	13	34	26	**215**
Department #3	10	11	13	40
Department #4	6	11	10	31
Department #5	5	20	11	19
Department #6	4	10	8	13
Department #7	3	8	8	9
Department #8	1	4	3	3
Department #9	1	-	-	-
Department #10	1	-	-	-

6.3 Phase Two Implementation and Results:

We use the software ILOG OPL for developing the IP model. Upon building the model, we first identify, in a rather arbitrary manner, the region for the parameter TP (i.e. the total "weekly patient throughput") within which global optimal OT plans exist. Specifically, the upper limit of the TP region is the one beyond which no more feasible global OT plan can be found; and the lower limit is one below which the same level of optimal global "weekly profit" is reported (i.e. when TP becomes a redundant constraint).

We then run the IP model on a variety of TP levels within the feasible region. The running results of the IP model on the various TP levels, each consisting of the total "weekly profit" generated and the allocation of the 98 OT slots to the 8 clinical departments are summarized in the Table 6.5 below.

Table 6.5: Global Optimal Pareto Set

Weekly Patient Throughput	Weekly Profit	Capacity Allocation							
		Dept #1	Dept #2	Dept #3	Dept #4	Dept #5	Dept #6	Dept #7	Dept #8
140	$120,165	22 (A)	28 (A)	7 (A)	8 (A)	18 (A)	7 (A)	6 (A)	2 (A)
141	$120,055	23 (B)	27 (B)	7 (A)	8 (A)	17 (B)	8 (B)	6 (A)	2 (A)
142	$119,990	22 (A)	28 (A)	8 (B)	8 (A)	17 (B)	7 (A)	6 (A)	2 (A)
143	$119,795	22 (C)	27 (B)	8 (B)	8 (A)	17 (B)	8 (B)	6 (A)	2 (A)
144	$119,615	22 (A)	27 (B)	9 (C)	8 (A)	17 (B)	7 (A)	6 (A)	2 (A)
145	$119,235	21 (D)	27 (B)	9 (C)	8 (A)	17 (B)	8 (B)	6 (A)	2 (A)
146	$118,650	20 (E)	27 (B)	10 (D)	8 (B)	17 (B)	8 (B)	6 (A)	2 (A)
147	$117,405	20 (E)	27 (C)	11 (E)	8 (A)	17 (B)	7 (C)	6 (A)	2 (A)
148	$115,865	21 (D)	28 (D)	11 (E)	8 (B)	13 (C)	8 (B)	7 (B)	2 (A)
149	$113,985	22 (F)	27 (C)	11 (E)	8 (A)	11 (D)	8 (B)	7 (B)	4 (B)
150	$110,880	23 (G)	25 (E)	11 (E)	9 (C)	11 (E)	8 (B)	7 (B)	4 (B)

In the table, the numbers stored in the entries under each department represent the total numbers of OT slots allocated to that department on the different patient throughput performance levels. The letters stored behind these numbers denote the exact candidate OT plans picked for each department, which will be explored further in the latter content. For example, in Department #1, there are altogether 4 capacity sizes (i.e. from 20 to 23 slots), and 7 (i.e. from "A" to "G") designs that are picked throughout the entire feasible TP region. It should also be noted that it is possible that two designs correspond to the same capacity level, e.g. both design A and design F in Department #1 consist of 22 slots, whilst they have different arrangements of the slots to the surgeons in the department.

Figure 6.3 plots the performances of the global Pareto solutions which are identified by the IP model.

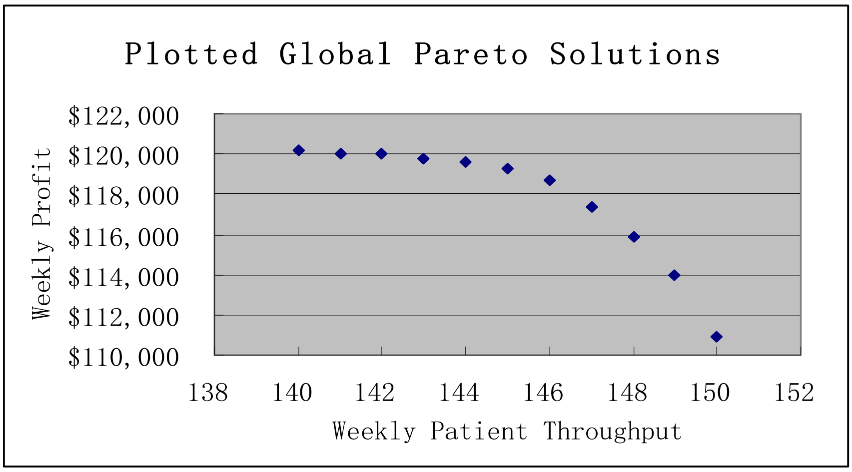

Figure 6.3: Plotted Global Pareto Solutions

It is revealed in Figure 6.3 that the objective values of the global Pareto solutions exist in a rather tight region (i.e. with the patient throughput ranging from 140 to 150 per week, and profit roughly from $ 110,000 to $120,000 per week). The reason for this, as we examined and later verified with the hospital, is that currently the total capacity at the OT is not really a serious constraint.

In fact, the narrow ranges of the Pareto solution performances inherently suggest that all the departmental OT plans picked in forming these global solutions have their own performances limited within narrow ranges. Recall our previous discussion, for any clinical department, it is only when the capacity allocated to it is increased to a sufficiently high level that the "gaps" between the performances of its Pareto solutions diminishes across the capacity levels, and that the performance scores entrapped into narrow regions. It can thus be implied that in running the IP model, all the departments have the luxury to receive sufficient OT slots. In other words, the total capacity at the OT is currently not a serious constraint.

After examining the performances of the global Pareto solutions, we further investigate the variations among these solutions in terms of their exact allocations of the OT slots to the surgeons.

In Table 6.5, it is shown that the slot allocation is rather robust within the global Pareto solutions on the “departmental” level. In particular, there are 4 out of 8 clinical departments, namely Departments #4, 6, 7, 8, which only have a variation of 1 or 2 slots in their allocated capacity among all their picked candidate OT plans. The Departments #1 and 2 both have a variation of 3 slots, and Department #3 4 slots. The only exception is Department #5, whose number of allocated slots varies between 11 and 18 slots.

Moreover, the following set of tables (Table 6.6A – Table 6.6C) depict, for three departments, the variations on the actual arrangements of the OT slots to the surgeons over the entire region of global optimal OT plans. As can be seen, only minor differences exist between the slots allocation between these picked designs. These are marked by the bold fonts in the table. In general, such differences are characterized either by the allocation of an additional dedicated slot to a surgeon who already has some, e.g. surgeons #1, 4, 16 in Departments #1 and surgeons #1, 5, 6, 10 in Department #2; or by a different choice of surgeon to leave the "shared group" and to receive a dedicated slot, e.g. surgeons #5 and 8 in Department #1 and surgeons #2, 5, 8, 10 in Departments #3.

Table 6.6A: Robustness of Departmental OT Plans Picked – Department #1

Design	Capacity	Slots Allocation																	
		Sur #1	Sur #2	Sur #3	Sur #4	Sur #5	Sur #6	Sur #7	Sur #8	Sur #9	Sur #10	Sur #11	Sur #12	Sur #13	Sur #14	Sur #15	Sur #16	Sur #17	Sur #18
A	22	3	3	2	3	.	.	1	.	.	2	2	.	.	1	2	1	.	1
B	23	3	3	2	3	.	.	1	.	.	2	2	.	.	1	2	2	.	1
C	22	3	3	2	2	.	.	1	.	.	2	2	.	.	1	2	2	.	1
D	21	2	3	2	2	.	.	1	.	.	2	2	.	.	1	2	2	.	1
E	20	2	3	2	2	.	.	1	.	.	2	2	.	.	1	2	1	.	1
F	22	2	3	2	2	.	.	1	1	.	2	2	.	.	1	2	2	.	1
G	23	2	3	2	2	1	.	1	1	.	2	2	.	.	1	2	2	.	1

Table 6.6B: Robustness of Departmental OT Plans Picked – Department #2

Design	Capacity	Slots Allocation												
		Sur #1	Sur #2	Sur #3	Sur #4	Sur #5	Sur #6	Sur #7	Sur #8	Sur #9	Sur #10	Sur #11	Sur #12	Sur #13
A	28	1	.	2	3	4	3	.	3	.	2	4	1	4
B	27	1	.	2	3	4	3	.	3	.	1	4	1	4
C	27	2	.	2	3	3	3	.	3	.	1	4	1	4
D	28	2	.	2	3	4	3	.	3	.	1	4	1	4
E	25	2	.	2	3	3	4	.	3	.	1	4	1	1

Table 6.6C: Robustness of Departmental OT Plans Picked – Department #3

Design	Capacity	Slots Allocation									
		Sur #1	Sur #2	Sur #3	Sur #4	Sur #5	Sur #6	Sur #7	Sur #8	Sur #9	Sur #10
A	7	3	.	.	.	.	.	1	1	1	.
B	8	3	.	.	.	.	1	1	1	1	.
C	9	3	.	.	.	.	1	1	1	1	1
D	10	3	1	.	.	.	1	1	1	1	1
E	11	3	1	.	.	1	2	1	.	1	1

We therefore conclude that the allocation of the OT slots, both on the departmental level and on the surgeons' level, are rather robust throughout the various global Pareto OT solutions.

Chapter 7: Conclusion

In this thesis, we have considered the capacity allocation problem at the operating theatre at a local hospital setting in Singapore. We have first proposed a new set of rules by which the future OT capacity allocations are done. According to these rules, the goodness of the capacity plans are evaluated by the “profit” and the “patient throughput” they generate for the hospital. Additionally, with respect to the fact that not all surgeons in the hospital can be provided with dedicated OT slots, we establish the rule of capacity sharing between surgeons from each department.

We develop a simulation model to mimic the dynamics of the surgery booking and execution processes, and to evaluate the performance of the OT capacity plans. In developing the model, we have made every effort to let it best represent the reality. Particular attention has been paid to issues covering the “actual demand” of each surgeon in the study period, the “queue-abandonment” behavior of the patients as well as the “surgery cancellation” patterns in each clinical department.

On top of the new rules set and the simulation model, we develop an algorithm that can effectively find the optimal allocations of the OT capacity for the hospital’s surgeons. Our algorithm exploits the special structure of the problem where patients are naturally grouped by their departments, and thus it tackles the overall capacity planning problem in two phases, i.e. the departmental and hospital phases. Our optimal solutions are expressed as Pareto sets, and thus freedom is allowed for the hospital administrators to pick any solution(s) to their best preference.

We claim that our new approach to allocating the OT capacity ensures the transparency and the fairness in this activity, as decision on all the surgeons' granted capacities are made at the central hospital management and by explicitly defined performance measures. In addition, our approach explicitly sets its objectives on the hospital's profitability and its total capability of serving the patients; therefore, if revisions on the capacity allocation can be done periodically through our approach, then both the financial and social goals of the hospital can be better achieved in the long run. Despite working on the local hospital's problem, we believe that our model and solution approach are general enough to be applied to OTs elsewhere.

One possible extension of our study may be to attempt to combine both the block-booking and FCFS systems at one OT setting. As we have discussed in Chapter 2, FCFS OTs ensure the resource utilization, keep the patients' waiting time at the minimal, yet lose the long-term vision on the system's performance; and block-booking system has exactly the opposite characters. Thus, it may be of the hospital's better interest if a hybrid OT utilization approach is implemented that strikes a balance between these two issues. As far as we see, there are generally two ways how one could practice both OT systems at one setting. The first way, obviously, is to make some ORs at the OT practice FCFS, and others practice block-booking; and the other way is to run the whole system generally under the block-booking framework, but set a cut-off time on the reserved booking and allow all unused capacity to be shared after this. In either way, there will be new decision variables introduced into the model (e.g. the "proportion of ORs to be run as FCFS", and/or where to set the cut-off time), and new performance measures that need to be evaluated that reflect the utilization and waiting time issues of the resource and the patients.

On the other hand, in our model all the resources surrounding the OT practice, e.g. the ICU beds, medical staff, and the surgeons themselves, are assumed to be available at all times. However, in actuality this might not be the case which renders our evaluation of the OT plans possibly overestimating the actual figures. Further study ought to be done on the characters of the resources availability, the mechanism how the system practice is modified when having unavailable resource, and consequently the implications of the resource availability on the system's overall performances, e.g. profit, patient throughput, etc.

Lastly, our model lends its concerns mostly with the "allocation" of the OT capacity. It would be an advancement if the "assignment" of the allocated slots to the surgeons can be tackled simultaneously in the model, while bearing with constraints addressing issues like the OR equipments' compatibility, surgeons' preferences, and so forth.

List of References

1. Blake, J.T., F. Dexter, and J. Donald. Operating room managers' use of integer programming for assigning block time to surgical groups: a case study. 2002. Anesthesia and Analgesia, v 94, p 143-148.
2. Blake, J.T., and J. Donald. Mount Sinai Hospital uses integer programming to allocate operating room time. 2002. Interfaces, v 32, n 2, p 63-73.
3. Dexter, F. and R.D. Traub. How to schedule elective surgical cases into specific operating rooms to maximize the efficiency of use of operating room time. 2002. Anesthesia and Analgesia, v 94, p 933-942.
4. Dexter, F., R.D. Traub, and P. Lebowitz. Scheduling a delay between different surgeons' cases in the same operating room on the same day using upper prediction bounds for case durations. 2001. Anesthesia and Analgesia, v 92, p 943-946.
5. Guinet, A. and S. Chaabane. Operating theatre planning. 2003. International Journal of Production Economics, v 85, n 1, p 69-81.
6. Hans, E., G. Wullink, M. van Houdenhoven, and G. Kazemier. Robust surgery loading. 2008. European Journal of Operational Research, v 185, p1038-1050.
7. Higgins, M. 2005. "Public hospitals decline swiftly", The Washington Times (August 17, 2005).
8. Hsu, V.N., R. de Matta, and C.Y. Lee. Scheduling patients in an ambulatory surgical center. 2003. Naval Research Logistics, v 50, n 3, p 218-238.
9. Jebali, A., A. B. Hadj Alouane, and P. Ladet. Operating rooms scheduling. 2006. International Journal of Production Economics, v 99, n 1-2, p 52-62.
10. Kuo, P.C., R.A. Schroeder, S. Mahaffey, and R.R. Bollinger. 2003. Journal of the American College of Surgeons, v 197, p 889-895.
11. Lamiri, M., X.L. Xie, A. Dolgui, and F. Grimaud. A stochastic model for operating room planning with elective and emergency demand for surgery. 2008. European Journal of Operational Research, v 185, p 1026-1037.
12. Lovejoy, W.S. and Y. Li. Hospital operating room capacity expansion. 2002. Management Science, v 48, n 11, p 1369-1387.
13. Marcon, E., S. Kharraja, and G. Simmonet. The operating theatre scheduling: an approach centered on the follow-up of the risk of no realization of the planning. 2001. Proceeding of the Industrial Engineering and Production Management, Quebec City, Canada.
14. Marcon, E., S. Kharraja, and G. Simonnet. Minimization of the risk of no realization for the planning of the surgical interventions into the operating theatre. 2001. IEEE Symposium on Emerging Technologies and Factory Automation, ETFA, v 1, p 675-680.
15. Ozkarahan, I. Allocation of surgeries to operating rooms by goal programming. 2000. Journal of Medical Systems, v 24, n 6, p339-378.

16. Singapore National Health Group. 2005. National Healthcare Group Annual Report 2004/2005.
17. Strum D.P, A.R. Sampson, J.H. May, L.G. Vargas. 2000a. Surgeon and type of anesthesia predict variability in surgical procedure times. Anesthesiology, v 92, p1454-1466.
18. Strum D.P., J.H. May, L.G. Vargas. 2000b. Modeling the uncertainty of surgical procedure times: comparison of the log-normal and normal models. Anesthesiology, v 92, p1160-1167.
19. United Nations Population Division. 2002. World population ageing 1950-2050.

Appendix A – Survey Results:

The survey results concerning the “Percentage of Lost Patients”, the “Queue Abandonment Behavior” (respectively for Private and Subsidized patients), and the “Percentage of Surgery Cancellations” in each clinical department are summarized as follows.

Department	Percentage of Lost Patients	Queue Abandonment (PTE Patients)	Queue Abandonment (SUB Patients)	Percentage of Cancellations
Department #1	8%	5 Weeks	6 Weeks	15%
Department #2	10%	7 Weeks	7 Weeks	10%
Department #3	5%	4 Weeks	4 Weeks	5%
Department #4	10%	7 Weeks	7 Weeks	15%
Department #5	8%	5 Weeks	6 Weeks	10%
Department #6	8%	5 Weeks	5 Weeks	10%
Department #7	5%	4 Weeks	4 Weeks	10%
Department #8	10%	6 Weeks	7 Weeks	20%

As shown, the percentage of lost patients due to the long waiting time is generally minimal as reported by all the eight clinical departments. This is in accordance with our conclusion that currently the total capacity is not a serious constraint at MOT. Moreover, private patients have shorter or the same tolerance level of waiting time in every clinical department as compared to the subsidized patients.

Appendix B – Summary of Surgeons' Information in Study Period:

Information concerning the "total profit", "total caseload" and "total surgery duration" of surgeons from all eight clinical departments in the study period are summarized in the tables below. In particular, surgeons corresponding to large amount of profit and/or large caseload are marked by bold fonts. As can be seen in Appendix C, these are the surgeons who are allocated most amount of OT capacity as the algorithm is executed.

Department #1:

Surgeon	Total Profit	Total Caseload	Total Duration
1	**$145,369**	**125**	**175**
2	**$123,319**	**161**	**223**
3	**$74,112**	**106**	**144**
4	**$121,616**	**104**	**195**
5	$19,840	36	45
6	$45,438	31	77
7	$36,446	74	89
8	$43,895	47	106
9	$20,042	42	29
10	**$155,212**	**88**	**175**
11	**$120,333**	**93**	**152**
12	$19,603	26	54
13	$15,181	23	36
14	$54,849	80	147
15	**$86,136**	**120**	**173**
16	$59,859	90	153
17	$27,451	35	59
18	$53,154	123	113

Department #2:

Surgeon	Total Profit	Total Caseload	Total Duration
1	$52,352	128	114
2	$14,112	22	29
3	**$87,204**	**96**	**139**
4	**$138,081**	**213**	**267**
5	**$199,279**	**131**	**346**
6	**$174,546**	**246**	**266**
7	$25,713	20	33
8	**$121,976**	**158**	**207**
9	$28,055	41	46
10	$66,776	63	109
11	**$178,404**	**187**	**302**
12	$50,108	61	109
13	**$208,350**	**100**	**299**

Department #3:

Surgeon	Total Profit	Total Caseload	Total Duration
1	**$155,052**	**175**	**187**
2	$1,933	88	52
3	$6,597	41	32
4	$13,906	45	40
5	$9,637	56	46
6	$42,714	113	150
7	$41,389	73	87
8	$35,857	42	34
9	$29,183	81	65
10	$15,556	93	75

Department #4:

Surgeon	Total Profit	Total Caseload	Total Duration
1	$42,900	67	85
2	$65,734	72	118
3	$34,194	154	106
4	**$182,699**	**151**	**250**
5	$24,270	40	29
6	**$116,053**	**116**	**171**

Department #5:

Surgeon	Total Profit	Total Caseload	Total Duration
1	**$151,110**	**74**	**232**
2	**$235,128**	**144**	**433**
3	**$136,220**	**158**	**248**
4	**$216,014**	**95**	**249**
5	**$252,622**	**163**	**486**

Department #6:

Surgeon	Total Profit	Total Caseload	Total Duration
1	**$111,766**	**133**	**203**
2	**$93,922**	**166**	**207**
3	$44,310	78	105
4	**$86,113**	**80**	**127**

Department #7:

Surgeon	Total Profit	Total Caseload	Total Duration
1	**$87,832**	**149**	**177**
2	**$80,961**	**168**	**187**
3	**$94,898**	**186**	**200**

Department #8:

Surgeon	Total Profit	Total Caseload	Total Duration
1	**$104,139**	**220**	**319**

Appendix C – Robustness of Departmental OT Plans Picked:

The following tables comprise the complete sets of candidate OT plans picked by the IP model for each clinical department. The arrangement of the OT slots to the surgeons is shown for each plan. As we have discussed, only minor differences exist between the slots allocation in each department's picked designs. These are marked by the bold fonts in the tables.

Department #1:

Design	Capacity	Slots Allocation																	
		Sur #1	Sur #2	Sur #3	**Sur #4**	**Sur #5**	Sur #6	Sur #7	**Sur #8**	Sur #9	Sur #10	Sur #11	Sur #12	Sur #13	Sur #14	Sur #15	**Sur #16**	Sur #17	Sur #18
A	22	**3**	3	2	**3**	**.**	.	1	**.**	.	2	2	.	.	1	2	**1**	.	1
B	23	**3**	3	2	**3**	**.**	.	1	**.**	.	2	2	.	.	1	2	**2**	.	1
C	22	**3**	3	2	**2**	**.**	.	1	**.**	.	2	2	.	.	1	2	**2**	.	1
D	21	**2**	3	2	**2**	**.**	.	1	**.**	.	2	2	.	.	1	2	**2**	.	1
E	20	**2**	3	2	**2**	**.**	.	1	**.**	.	2	2	.	.	1	2	**1**	.	1
F	22	**2**	3	2	**2**	**.**	.	1	**1**	.	2	2	.	.	1	2	**2**	.	1
G	23	**2**	3	2	**2**	**1**	.	1	**1**	.	2	2	.	.	1	2	**2**	.	1

Department #2:

Design	Capacity	Slots Allocation												
		Sur #1	Sur #2	Sur #3	Sur #4	**Sur #5**	**Sur #6**	Sur #7	Sur #8	Sur #9	**Sur #10**	Sur #11	Sur #12	**Sur #13**
A	28	**1**	.	2	3	**4**	**3**	.	3	.	**2**	4	1	**4**
B	27	**1**	.	2	3	**4**	**3**	.	3	.	**1**	4	1	**4**
C	27	**2**	.	2	3	**3**	**3**	.	3	.	**1**	4	1	**4**
D	28	**2**	.	2	3	**4**	**3**	.	3	.	**1**	4	1	**4**
E	25	**2**	.	2	3	**3**	**4**	.	3	.	**1**	4	1	**1**

Department #3:

Design	Capacity	Slots Allocation									
		Sur #1	Sur #2	Sur #3	Sur #4	Sur #5	Sur #6	Sur #7	Sur #8	Sur #9	Sur #10
A	7	3	.	.	.	.	.	1	1	1	.
B	8	3	.	.	.	.	1	1	1	1	.
C	9	3	.	.	.	.	1	1	1	1	1
D	10	3	1	.	.	.	1	1	1	1	1
E	11	3	1	.	.	1	2	1	.	1	1

Department #4:

Design	Capacity	Slots Allocation					
		Sur #1	Sur #2	Sur #3	Sur #4	Sur #5	Sur #6
A	8	.	1	1	3	.	2
B	8	1	.	1	3	.	2
C	9	.	1	2	3	.	2

Department #5:

Design	Capacity	Slots Allocation				
		Sur #1	Sur #2	Sur #3	Sur #4	Sur #5
A	18	3	4	3	4	4
B	17	3	4	3	3	4
C	13	1	3	3	3	3
D	11	1	1	3	3	3
E	11	1	3	3	3	1

Department #6:

Design	Capacity	Slots Allocation			
		Sur #1	Sur #2	Sur #3	Sur #4
A	7	2	2	1	2
B	8	2	3	1	2
C	7	2	3	1	1

Department #7:

Design	Capacity	Slots Allocation		
		Sur #1	Sur #2	Sur #3
A	6	2	2	2
B	7	2	2	3

Department #8:

Design	Capacity	Slots Allocation
		Sur #1
A	2	2
B	4	4